UNIVERSITY OF CALIFORNIA,
IRVINE

Improving Efficacy of Support Groups in Online Environments

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Management

by

Ali Esmaeeli

Dissertation Committee:
Professor L. Robin Keller, Co-chair
Professor Cornelia (Connie) Pechmann, Co-chair
Associate Professor John Turner

2020

DEDICATION

To

My supportive family who value diligence and wisdom

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I would like to express my deepest appreciation to my advisors, Professor L. Robin Keller and Professor Cornelia (Connie) Pechmann, for their unwavering support and guidance throughout this entire process. Thanks also to Professor John Turner for his valuable insights and suggestions that helped me in my research.

I am extremely grateful to my wife, Marzieh, for being a great companion throughout this journey. My success would not have been possible without her support.

This research was supported by National Cancer Institute Grant R01CA204356-04.

VITA

Ali Esmaeeli

2007	B.S. in Information Technology Engineering, Amirkabir University of Technology
2010	M.S. in Information Technology Engineering, Amirkabir University of Technology
2020	Ph.D. in Management, University of California, Irvine

PUBLICATIONS

A. Esmaeeli, C. Pechmann, and J.J. Prochaska, Examination of a Homophily-based Buddy System for Online Support Groups: Relationships with Tie Strength and Goal Attainment, Revision requested at *Journal of Interactive Marketing*

M. Razzazi and A. Esmaeeli, Balance Allocation Mechanism: An Optimal Mechanism for Multiple Keywords Sponsored Search Auctions, *Information Sciences*, vol. 262, March 2014 http://dx.doi.org/10.1016/j.ins.2013.08.033, pp. 190-214

A. Esmaeeli and H.R. Shahriari, Privacy Protection of Grid Service Requesters through Distributed Attribute Based Access Control Model, Proc. of 5th Int. Conf. on Advances in Grid and Pervasive Computing, Lecture Notes in Computer Science, Springer-Verlag, Taiwan, May 2010, pp. 573-582

A. Esmaeeli and M. Shajari, MVPayword: Secure and Efficient Payword-Based Micropayment Scheme, Proc. of 2nd Int. Conf. on the Applications of Digital Information and Web Technologies, London, Aug. 2009, pp. 609-614

ABSTRACT OF THE DISSERTATION

Improving Efficacy of Support Groups in Online Environments

by

Ali Esmaeeli

Doctor of Philosophy in Management

University of California, Irvine, 2020

Professor L. Robin Keller and Professor Cornelia (Connie) Pechmann, Co-chairs

In this research I evaluated two strategies for improving performance of online support groups. The first strategy was to use a buddy system in online support groups. This system involves pairing demographically similar members of a group to serve as each other's buddies. Analyses of real online support groups indicated that members whose only difference with other members is in having a buddy who is more active rather than less active in the group are more engaged in the group and have a higher chance of attaining their goal for which they joined the group. Analysis also showed that members help their buddies more than they help other group members. Results are robust across four measures of tie strength, including contact frequency, reciprocity, and two measures of contact length. Overall, results suggest that managers can use the buddy system in online support groups as an effective method to drive group engagement and increase the support provided and goal attainment.

The second strategy was to add a chatbot to the online support groups to provide members with additional informational and motivational support. The design of the chatbot which can respond to members' messages based on their content is presented here.

The chatbot has a natural language understanding component which can identify intent of messages out of 26 possible intents related to smoking cessation which is the main context of the online support groups studied here. The chatbot is supposed to respond to messages with 25 intents called triggering intents and ignore 1 intent which is non-triggering intent. Triggering intents account for less than half of the total messages. The bot is intentionally designed to have higher precision than recall for triggering intents. Precision for triggering intents is the probability of identifying the intent correctly if the message is identified as a triggering intent. Recall for triggering intents is the percentage of triggering messages which are identified with correct intent. The chatbot has the precision of 81% and the recall of 44% for triggering intents. Separate randomized control trial experiments are required to evaluate the overall performance of the chatbot in online support groups.

INTRODUCTION

Online social networks are very popular, accounting for about one third of the time that people spend online (Buckle 2016). Users spend more than two hours per day on average on online social networks (Mander 2017). This availability and popularity has created a framework for new and innovative applications of online social networks. Examples of these applications are online support groups for health behavior change (Barrera et al. 2002; Cavallo et al. 2012; Maher et al. 2014; Pechmann et al. 2017; Setoyama, Yamazaki and Namayama 2011; Turner-McGrievy and Tate 2013) and collaborative learning (Cho et al. 2007; Wasko and Faraj 2005). Online support groups seek to provide members with social support so that they achieve their goals. These groups attract people with specific goals, interests, or needs.

Online support groups are also important for firms. Firms may benefit from online support groups by gaining access to market research information (Bickart and Schindler 2001; Kozinets 2002; Divakaran et al. 2017; Gruner, Homburg and Lukas 2014), charging for the online services offered (Armstrong and Hagel III 1996), direct selling (Naylor, Lamberton and West 2012), advertising to the members (Armstrong and Hagel III 1996), and word-of-mouth promotions (Brown, Broderick and Lee 2007).

The main application of online support groups that I focused my research on is in the field of health behavior change. Online support groups seek to provide peer social support for health behavior change by ensuring positive and lasting interactions between the online group members (Moorhead et al. 2013; Portnoy et al. 2008). Examples of these applications include smoking cessation (Cobb, Graham and Abrams 2010; Graham et al. 2017; Pechmann et al. 2015; Pechmann et al. 2017), diabetes control (Barrera et al. 2002), breast cancer support (Setoyama, Yamazaki and Namayama 2011), weight loss (Bradford, Grier and Henderson 2017; Brindal et al. 2012; Hwang et al. 2010; Napolitano et al. 2013; Turner-McGrievy and Tate 2013; Parkinson et al. 2017), physical activity promotion (Cavallo et al. 2012; Foster et al. 2010; Valle et al. 2013), and health information provision (Kuwata et al. 2010; Freyne et al. 2010).

Collaborative learning is another important domain that benefits from online social networks. Learning involves the process of knowledge acquisition from social relations and communications with other people (Brown and Duguid 1991). Collaborative learning communities provide knowledge sharing as well as social support for educational attainment (Cho et al. 2007; Haythornthwaite 2002a; Wasko and Faraj 2005). Engagement takes place through computer-supported collaborative learning (Cho et al. 2007) and online discussion forums (Wasko and Faraj 2005). Social interaction is a key element in these collaborative learning environments (Kreijns, Kirschner and Jochems 2003).

The basic requirement for success in online groups is to have active members who provide social support to each other (Arguello et al. 2006). Social media companies like Snapchat (Snap Inc. 2019) work hard to add features such as Snapstreaks to increase social interactions in their online networks. Many studies have also noted the importance of member engagement and interactions in online groups (Maher et al. 2014). Ensuring active engagement and interactions between online group members have been shown to be important for online weight loss programs (Turner-McGrievy and Tate 2013), online smoking cessation programs (Cobb et al. 2005), collaborative learning programs (Kreijns, Kirschner and Jochems 2003), and breast cancer support communities (Setoyama, Yamazaki and Namayama 2011). Therefore, to have effective online support groups, methods are required to increase member engagement and social interactions with each other. Those who design online support groups seek to provide members with tools that will help them to contribute and provide social support to their groups (Arguello et al. 2006). In this work, I study two methods for improving the efficacy of online support groups for members. In Chapter 1, I look at the effect of adding a buddy system on goal attainment and also on interactions between members in online support groups. In Chapter 2, I present the design of a chatbot with natural language processing capabilities that can be used to improve the performance of online support groups by automatically responding to members' messages and providing them with informational and motivational support.

CHAPTER 1: The Effects of Buddy Systems in Online Support Groups on Members' Goal Attainment and Interactions[1]

Buddy systems are increasingly prevalent in online communities. In a typical buddy system, members are provided with peer support from a buddy to help them achieve their goal. There are several applications for online buddy systems in the areas of health, physical activity, and behavior change (The Monday Campaigns 2019; Broc & Bells 2019; Workout Buddies 2019). However, most online buddy systems have not been evaluated. Therefore, despite the widespread use of online buddy systems, there is a lack of scientific research that addresses the efficacy and dynamics of these systems.

Many researchers have studied buddy systems in face-to-face or offline contexts. A 2002 national survey of 7467 U.S. mental health support groups found that 38% of them provided face-to-face buddy systems in addition to general support for the members (Goldstrom et al. 2006). A systematic review of cancer peer-support studies found that 30% (13 of 43) had implemented buddy systems, either in person or over the phone (Hoey et al. 2008). Two separate literature reviews on buddy systems for people with cancer found 28 studies published from 1966-2007 and 13 studies published from 2007-2014 (Macvean, White and Sanson-Fisher 2008; Meyer, Coroiu and Korner 2015). Beyond health, offline buddy systems have been studied for crime control (Fo and O'Donnell 1975;

---

[1] I presented different versions of materials in this chapter at 2017 INFORMS Annual Meeting Houston with the title of "The Effects Of Assigning Buddies In Online Health Communities", and 2018 INFORMS Annual Meeting Phoenix with the title of "The Effects of Assigning Buddies on Tie Strength in Online Communities". Also, a paper co-authored by Professor Cornelia Pechmann (cpechman@uci.edu) and Professor Judith Prochaska (jpro@stanford.edu) with the title of "Examination of a Homophily-based Buddy System for Online Support Groups: Relationships with Tie Strength and Goal Attainment" was submitted to the Journal of Interactive Marketing and is requested for a revision.

O'Donnell and Williams 2013), learning and professional development (Guhde 2005; Kukulska-Hulme and Pettit 2008), and improving children's social skills (English et al. 1997; Hektner et al. 2017).

There is some evidence that buddy systems work well as part of face-to-face or offline services (West, Edwards and Hajek 1998; Lee et al. 2013; Nicholas and Keilty 2007; Zuyderduin, Ehlers and Van der Wal 2008). However, it is not clear that buddy systems will provide the necessary social support to their partners in online service contexts. There is considerable evidence that people's engagement in online communities is often low (Preece, Nonnecke and Andrews 2004; Beenen et al. 2004). One study found that the rate of lurkers (those who do not contribute enough to the community) in health support communities is around 46% and for software support communities this rate is about 82% (Nonnecke and Preece 2000). Moreover, studies show that relationships originated in online environments are less close and supportive compared to face-to-face ones (Mesch and Talmud 2006; Trepte, Dienlin and Reinecke 2015). The online realm lacks standard demographic cues that can help people get to know each other. People who are online often resist self-disclosing personal information which would allow their buddies to get to know them; or people who are online may even falsify their personal information (Belk 2013; Nguyen, Bin and Campbell 2012).

These potential problems with online buddy systems, and the dearth of research on them, motivated me to implement and evaluate the buddy system for online support groups to see whether relatively active buddies actually help their partners. I studied a buddy system that was being used in an online smoking cessation program. I addressed the following research questions: 1. In an online buddy system, will active buddies emerge who facilitate goal attainment for their partners? 2. In an online buddy system, will the buddies do their job and form stronger ties with their partners, i.e., post messages more to their partners, rather than posting to other group members instead? 3. In an online buddy system, will strong buddy ties serve to strengthen ties with others as well, further facilitating goal attainment through enhanced social support? 4. In an online buddy system, will strong ties from the buddy to the partner, and then the partner to others, mediate effects on abstinence from smoking of the partner?

Consistent with the social network literature (Friedkin 1980; Gilbert and Karahalios 2009; Granovetter 1973; Marsden and Campbell 1984; Petróczi, Nepusz and Bazsó 2007), I used four measures to evaluate ties between dyadic members in online support groups, including one measure of interaction frequency, one measure of reciprocity, and two measures of time spent interacting (contact time and duration of contact).

My research differs from previous studies on offline buddy systems in three main ways. First, I looked at a complete record of members' behavior with their buddies and others in the social network to evaluate social support, instead of asking members to fill out self-report questionnaires (Gruder et al. 1993; Hennrikus et al. 2010). This is the advantage I got from the online environment. Second, I compared the peer support that the buddy provided to the partner with the peer support that he/she provided to other people in the group, instead of looking solely at the buddy-partner relationship (Gruder et al. 1993; Hennrikus et al. 2010). Third, in my research, the only communication medium between the members was the informal social network that was provided online; whereas in previous studies of offline buddies, the members communicated in both informal social networks and formal groups that were run by experts (May et al. 2006b; Gruder et al. 1993).

Theoretical Framework and Hypotheses

Buddy Systems and Hypotheses Development

In a broad definition, a buddy system is an arrangement in which two individuals are paired (Merriam-Webster Inc. 2019). In a typical buddy system, individuals provide social support to their buddies to help them achieve their goal more easily. Online buddy systems are very popular, especially in health and learning contexts. For example, BuddySystem is a project supported by the European University Foundation to help international students find local buddies that help them with different needs during their transition to the new environment (Buddy System 2019). Quit and Stay Quit Monday which is a program for smoking cessation specifies finding a quit buddy as one of its main guidelines that can help quitters (The Monday Campaigns 2019). There are several online platforms and cellphone applications that help members to find buddies for different health and well-being activities such as fitness and exercise (Broc & Bells 2019; Workout Buddies 2019). There are also several online frameworks for providing cancer fighters, survivors, and caregivers with buddy support from similar people (PLWC 2019; Imerman Angels 2019). One study of a mobile phone buddy system intervention for women with diabetes in South Africa found short-term increases in positive action and coping, but increases in blood glucose and diastolic blood pressure (Rotheram-Borus et al. 2012). Another study that evaluated the effects of a nonstandard buddy system in which members could select a different person as their buddies each day observed negative outcomes perhaps due to reactance (Kim and Sundar 2014). However, most online buddy systems are not evaluated. While there are different examples of implemented online buddy systems designed to help people achieve their goals easier, there is lack of scientific research on them. I am especially interested in the efficacy of these systems and their dynamics in online support groups.

In contrast to the online setting, for offline and face-to-face contexts, buddy systems have been used and studied for several years. A 2002 national survey of 7467 mental health groups containing mutual support groups, self-help organizations, and consumer-operated services in the U.S. found that 38% of these groups provided a face-to-face buddy system in addition to mutual support for their members, including 25% of mutual support groups, 43% of self-help organizations, and 60% of consumer-operated services (Goldstrom et al. 2006). Buddy systems are also common in peer support groups; a systematic review of cancer peer-support programs found that 13 of 43 studies implemented buddy systems, either in person or over the phone (Hoey et al. 2008). Separate literature reviews on programs with offline buddy systems for people with cancer found 28 papers published from 1966-2007 and 13 papers published from 2007-2014 (Macvean, White and Sanson-Fisher 2008; Meyer, Coroiu and Korner 2015).

Table 1.1 shows specific examples of applications of offline buddy systems. Many studies have already shown the positive effects of these systems. Dyadic peer support intervention in newly diagnosed breast cancer patients in Korea improved self-efficacy compared to usual care (Lee et al. 2013). A parent-to-parent peer support intervention matched parents of medical technology-assisted children with chronic lung disease who had similar caregiving responsibilities. Researchers found sharing daily experiences helped reduce isolation, increase knowledge, and provide an important sense of feeling understood (Nicholas and Keilty 2007). Smokers attempting to quit with a buddy are significantly more likely to achieve abstinence and stay abstinent than those attempting to quit alone (West, Edwards and Hajek 1998). HIV+ women in Botswana who were paired with a peer buddy also showed increased self-care behaviors (Zuyderduin, Ehlers and Van der Wal 2008). Youths paired with adult buddies are less likely to commit crimes, even 35 years after the mentoring program ended (Fo and O'Donnell 1975; O'Donnell and Williams 2013).There are more studies that show the positive effects of offline buddy systems in smoking cessation (Gruder et al. 1993; Kviz et al. 1994; Murray et al. 1995), professional development (Guhde 2005), children's social skills development (Hektner, Brennan and August 2017), helping students with autism (McCurdy and Cole 2014), and alcohol use reduction (Tevyaw et al. 2007).

Table 1.1 Studies of offline buddy systems

People Serviced	Buddy System Study
Smokers	Albrecht et al. (2006); Carlson et al. (2002); Donatelle et al. (2000); Gruder et al. (1993); Hennrikus et al. (2010); Kviz et al. (1994); May et al. (2006); Murray at al. (1995); West, Edwards and Hajek (1998)
Alcoholics	Fals-Stewart, Birchler and Kelley (2006); Tevyaw et al. (2007)
Dieters	Morgan et al. (2011); Napolitano et al. (2013)
Diabetics	Keogh et al. 2011; Rotheram-Borus et al. (2012)
Breast cancer victims	Lee et al. (2013)
HIV/AIDS victims	Zuyderduin, Ehlers and Van der Wal (2008)
Parents with chronically ill children	Nicholas and Keilty (2007)
Adults needing professional skills	Guhde (2005); Kukulska-Hulme and Pettit (2008)
Children needing social skills	English et al. (1997); Hektner, Brennan and August (2017)
Children with autism	McCurdy and Cole (2014)
Delinquent children	Fo and O'Donnell (1975); O'Donnell and Williams (2013)

Although there is some established evidence showing the efficacy of buddy systems in the offline context, it is not clear that similar effects will exist for buddy systems in an online context. There are substantial differences between face to face and online contexts. Studies show that people's engagement in online communities is often low (Preece, Nonnecke and Andrews 2004; Beenen et al. 2004). One study shows that the percentage of lurkers in software support communities is as high as 82% and this rate is 46% for health support communities (Nonnecke and Preece 2000). Relationships originated in online environments are less close and supportive than face-to-face ones (Mesch and Talmud 2006; Trepte, Dienlin and Reinecke 2015). Other differentiating features of online social exchanges include fewer individuating cues that distinguish conversation partners, more idealized perceptions of conversation partners, more control over self-disclosures, disinhibition in self-disclosures, and multiple online selves (Belk 2013; Nguyen, Bin and Campbell 2012).

The potential problems of online buddy systems, and the lack of scientific research on them, motivated me to implement and evaluate a buddy system for online support groups. I was specifically interested to see whether buddies who are more active as compared to those being less active actually help their partners and, if yes, how they do that.

My first hypothesis addresses the first research question about the overall effect of buddy activity on partners' goal attainment in online support groups. This hypothesis predicts that the buddy system helps support group members whose buddies are more active to achieve their goal better than those with less active buddies. Buddy activity is measured as the average number of posts to the group per day.

H1. When buddies are more active in online support groups, their partners are more likely to achieve their goal.

Buddy systems can be useful if they add to the level of social support which is already provided to individuals within a support group. The efficacy of both offline and online support groups has been shown to depend on the amount of the social support that they provide to members (Hanson et al. 1990; Westmaas, Bontemps-Jones and Bauer 2010; Cobb et al. 2005). If a buddy system improves this social support, it can be useful in the support group. For example, for the application of smoking cessation, when there is minimal social contact and adding a buddy provides an extra level of social support for people, it shows a positive effect on goal attainment (Gruder et al. 1993; Kviz et al. 1994; West, Edwards and Hajek 1998). On the other hand, when there already exists a high level of social support, a buddy system does not show a significant positive effect (May et al. 2006b).

To see if the buddy system adds to the level of social support in online support groups, I needed to evaluate the support that buddies who are more active provide to their partners. If I observe the positive effect of the buddies' activity in the group on their partners' goal attainment, I also expect to see stronger support from these buddies to their partners in the groups. My second hypothesis evaluates this support.

To evaluate the support provided by the buddy in an online support group, I needed to measure ties between individuals in online groups. Ties between dyads is the most fundamental component of both online and offline social networks, and tie strength is the main outcome used to assess engagement with others and online group success (Aral and Walker 2014; Garton, Haythornthwaite and Wellman 1997; Haythornthwaite 2002b).

H2. When buddies are more active in online support groups, they form stronger ties with their partners.

To answer the second research question, in addition to H2, I also needed to compare the relationship that buddies form with their partners with the relationships that they form with other people in the group. In an effective online buddy system, I expect to see stronger relationships with partners compared to relationships with other people in the group. Hypothesis H3 evaluates this comparison.

H3. Buddies in online support groups form stronger ties with their partners than with other group members on average.

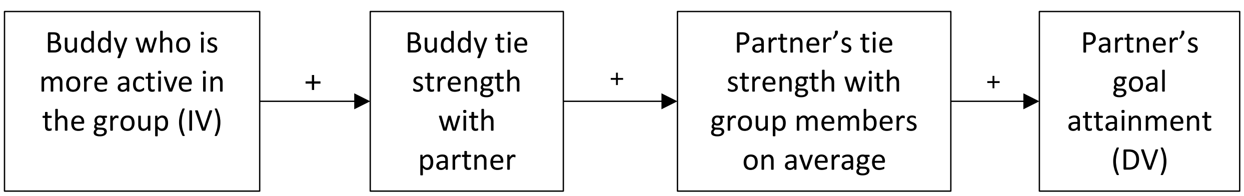
The support for this hypothesis suggests that the tie to the partner is especially strong, showing that buddies help the partner more than they help other people in the group.

The perception of online group members regarding the level of social support provided by their group is positively related to their engagement level with others in the group. For example, Obst and Stafurik (2010) show that perceived online social support is correlated with the amount of time spent communicating online with others. Therefore, a perceived higher level of social support by members (through stronger ties from buddies who are more active) is expected to improve interactions with all people in the group. Some studies also suggest that assigning online buddies in online communities may improve attachment to the group for people who are shy about posting and do not like the group (Du 2006; Preece, Nonnecke and Andrews 2004). Testing H4 provides the answer to the third research question.

H4. When buddies are more active in online support groups, their partners form stronger ties with group members on average.

Hypotheses H1, H2, and H4 are different parts of a casual model which shows how the buddy system works in online support groups. Buddies who are more active in a group form stronger ties to their partners. This extra level of support positively affects the interaction of partners with other members in the group and this subsequently increases the chance of goal attainment for partners (see Figure 1.1). The relation between different parts of this causal model is tested in a mediation test described in H5. H5 provides the answer to the fourth research question.

Figure 1.1 The effect of buddy activity on interactions and goal attainment



H5. When buddies are more active in online support groups, then partners are more likely to achieve their goal because a) the buddies form stronger ties with them, and in turn, b) they form stronger ties with group members on average. Support for this hypothesis explains how having a buddy who is more active in the group increases the chance of goal attainment (abstinence in this context) in online support groups.

Measuring Tie Strength

Tie strength refers to the strength of a dyadic social relationship based on the amount of time two people spend interacting and other important characteristics of their interactions including frequency of interactions and reciprocity (Granovetter 1973). Tie strength is important, e.g., political mobilization messages on Facebook have greater impact if network ties are strong (Bond et al. 2012). In addition, group recommender systems improve in performance if they include tie strength measures (Quijano-Sánchez, Díaz-Agudo and Recio-García 2014). Many researchers have devoted considerable time and effort to measuring tie strength in both online and offline settings (Friedkin 1980; Gilbert and Karahalios 2009; Marsden and Campbell 1984; Petróczi, Nepusz and Bazsó 2007) and these works provide the foundation for the tie measurement approach presented here.

Frequency is the most commonly measured component of tie strength. It usually refers to the frequency of contact between dyads in a relationship (Jeners, Nicolaescu and Prinz 2012; Wiese et al. 2015). Examples of variables that are used to evaluate frequency are ordinal variables with levels like "once a year" or "more than once a week" (Cummings, Butler and Kraut 2002; Granovetter 1973; Marsden and Campbell 1984; Reagans 2011) and continuous variables like number of posts, comments, or likes in social media (Arnaboldi, Guazzini and Passarella 2013; Burke and Kraut 2014; Gilbert, Karahalios and Sandvig 2008; Gilbert and Karahalios 2009; Jones et al. 2013; Panovich, Miller and Karger 2012). I measure the frequency of messages that a person sends to another person as one of the tie strength measures.

Reciprocity is another key component of tie strength (Granovetter 1973; Friedkin 1980; Plickert, Cote and Wellman 2007). Strong ties between dyads are those that are reciprocated (Friedkin 1980; Petróczi, Nepusz and Bazsó 2007). In previous studies, different measures were used to evaluate reciprocity in a relationship including whether comments were exchanged on social media (Gilbert, Karahalios and Sandvig 2008), whether applications were shared (Gilbert and Karahalios 2009), and whether interactions were non-directional, unidirectional, or bidirectional (Jeners, Nicolaescu and Prinz 2012). In this research, I measure the maximum number of consecutive days with bidirectional interactions as the tie strength reciprocity measure.

Other important measures of tie strength relate to the amount of time two people spend interacting. This is often assessed based on total contact time, e.g., the total number of days where there was contact, i.e., excluding non-contact days (Granovetter 1973; Marsden and Campbell 1984; Mathews et al. 1998). Other studies also consider the duration of the relationship, by comparing the first and last days (or years) when people are in contact, which differs from total contact time because it includes non-contact days (Arnaboldi, Guazzini and Passarella 2013; Gilbert and Karahalios 2009; Marsden and Campbell 1984; Panovich, Miller and Karger 2012). I measure both total contact time and relationship duration as tie strength measures.

These four tie strength measures share three main characteristics. The first characteristic relates to the way that I collect the data. In most offline settings (Marsden and Campbell 1984) and even in some online settings (Liberatore and Quijano-Sanchez 2017; Petróczi, Nepusz and Bazsó 2007), questionnaires are used to calculate the values of the variables that are used to estimate tie strength. More recent research uses social media data to estimate tie strength (Arnaboldi, Guazzini and Passarella 2013; Gilbert and Karahalios 2009; Xiang, Neville and Rogati 2010). Data mining techniques are also used by some researchers to predict tie strength from social media data (Sohrabi and Akbari 2016). I use the complete record of members' behavior in a closed online group to measure tie strength. The second characteristic is related to the precision of the measurement of tie strength. Earlier studies that evaluated tie strength in offline social networks usually partitioned relations into weak and strong ties (Granovetter 1973; Marsden and Campbell 1984). In this study, like in many recent studies (Arnaboldi, Guazzini and Passarella 2013; Gilbert and Karahalios 2009; Petróczi, Nepusz and Bazsó 2007), I use more precise continuous or interval measures of tie strength.

The third characteristic of the measurement approach presented here is related to how I address relationship symmetry. Many previous studies assumed that the relations between people involved symmetric ties, and so they asked just one person in the dyad questions about tie strength (Gilbert and Karahalios 2009; Marsden and Campbell 1984). This assumption may be reasonable when ties are assigned dichotomous values such as positive/negative or weak/strong. When a quantitative continuous measurement scale is employed, many ties are not symmetric anymore. Even when both sides of a tie are active members, the strength of ties can differ. Like Petróczi, Nepusz and Bazsó (2007), I consider asymmetric relations when I calculate tie strength, except for the reciprocity measure which is symmetric in nature.

Research Methods

Setting

I conduct analyses using the data from the Tweet2Quit project. This project is a real treatment program funded by NIH for designing effective Twitter-based online support groups for quitting smoking. Participants were selected based on the requirements listed in the research protocol that is posted on clinicaltrials.gov. Selected participants were organized into online groups of 20 members which is the size of social network for an average person (Trusov, Bodapati and Bucklin 2010). Other studies have also used a similar number of members in their support groups both for online (Napolitano et al. 2013) or offline contexts (May et al. 2006a; Tevyaw et al. 2007). 36 online support groups were studied from 2016 to 2019 (N = 720). Within each group, the members were added at the same time, and they followed each other on Twitter for 90 days. Also, the groups were set up to be private, in that no outside members could follow or be followed by the group members (similar to a standard GroupMe setup). Therefore, the basic structure of each online group did not change during the 90 days. The only communication medium between group members was tweeting within the group. When a member posted within the group, all other members could see the posts. The demographics of members in a group were not disclosed to other members of the group. Unless they self-disclose their information in their messages, they cannot develop relationships with others based on similarity in demographics.

When organizing each support group, pairs of group members were assigned to serve as each other's buddy, so there were 10 buddy pairs in each 20-member group. Buddies were assigned based on being the same gender (male or female); and then based on similarity in age, geographic location (according to zip code), and years of education, with equal weight assigned to each of these three factors. Members were informed of their buddies via email, when their online groups started, but members were not told how their buddies were selected. Members communicated with their buddies in the same way that they communicated with the non-buddy members. They were advised to develop good relations with their buddies. However, they could choose not to do so.

The length of the program for each group was 90 days. All group members were advised to quit smoking during the first 10 days. To have data about goal attainment in each group, abstinence information was collected from members after one month (initial abstinence) and after 3 months (maintained abstinence) in the program. Each member received free FDA-approved nicotine replacement therapy (NRT) patches which last for at least 8 weeks to help them with quitting.

When online group members tweet, they can use “@username” to mention other specific members of the community. On Twitter, replying to a tweet automatically creates a mention, i.e., it adds @username. Tweets involving mentions are viewable to the entire online group, but the intended recipient who is mentioned may be notified of the mention, if this type of notification has been activated on the recipient’s mobile or other device. It has been estimated that 25.4% of all Twitter posts mention someone (Huberman, Romero and Wu 2009). Mentions are an important sign of existing relationships between Twitter users (Huberman, Romero and Wu 2009). In this study, mentions were used to evaluate the tie strength between online support group members. When member m_1 mentions member m_2 in a tweet, this is considered as a post from m_1 to m_2 ($m_1 \rightarrow m_2$). Note that members may mention more than one other member in a tweet. For example, m_1 may mention both m_2 and m_3 in the same tweet. This is counted as one mention from m_1 to m_2, and one mention from m_1 to m_3 (like when you email the same message to several people).

Dependent and Independent Variables

Abstinence Status The main dependent variable in this research is abstinence status, which measures a member’s goal attainment. It is used to evaluate how successful the online support group members are in achieving their goal. Using surveys, 7-day abstinence data were gathered from each group member at the end of the first month and third month. For the analyses in this chapter, the first thirty days are considered as time frame 1 and the next sixty days are considered as time frame 2. Time frames are designed to not be equal in length so that they represent different phases in the quitting process. If members smoke during the last seven days of the time frame when they fill out the survey, they are considered as non-abstinent. They can still use NRT patches (which are part of the study) and be counted as abstinent from smoking. Abstinence data is used as the dependent variable for hypotheses H1 and H5.

Buddy Activity The overall activity of a buddy in the group is the main independent variable in the analyses of this chapter. This variable is defined for each member during a time frame and is measured in terms of the average number of posts to the group per day. In hypotheses, the effect of a buddy's activity level on the partner's goal attainment and the level of interactions with others in the group is evaluated.

Tie Strength Measures I used four quantitative and continuous tie strength measures to evaluate members' engagement with others during each time frame in online support groups. Analyzing tie strength provides a better understanding of the dynamics of buddy systems in online support groups, which made my results generalizable to other types of online support groups. The tie strength measures that I defined and used here are quite general and are applicable to various types of online communities. I used mentions of members in tweets to calculate each measure. For the remainder of this chapter, m_1 and m_2 refer to a member and one of his/her peers in an online support group respectively.

The first measure of tie strength is frequency per day. For a specific time frame, I evaluated the frequency per day of the relationship from member m_1 to member m_2 ($m_1 \rightarrow m_2$) by calculating the average number of times in a day that m_1 mentioned m_2 during that time frame. Frequency per day may be asymmetric, meaning that the frequency of the relationship $m_1 \rightarrow m_2$ may be different from the frequency of the relationship $m_2 \rightarrow m_1$. Since the lengths of time frames are not equal, I used a daily average instead of the total numbers. This enabled me to have the same scale for the tie strength frequency across different time frames.

The second measure of tie strength is the level of reciprocity days. For a specific time frame, the longest number of reciprocated days refers to the count of consecutive days that both members of the dyad, members m_1 and m_2, mentioned each other at least once. I used the longest number of reciprocated days, and then converted this to a percentage of the total number of days in the time frame, so that I have the same scale across time frames. This measure is similar to the Snapstreaks measure used by Snapchat. Since this measure evaluates reciprocity in a dyad, unlike the previous measure, it is symmetric by nature. This means that the value of reciprocity days for the relationship $m_1 \rightarrow m_2$ is equal to the value of reciprocity days for the relationship $m_2 \rightarrow m_1$.

The third measure of tie strength is contact days. For a specific time frame, I computed the contact days for the relationship $m_1 \rightarrow m_2$ by calculating the percentage of days that member m_1 mentioned member m_2 at least once. Like frequency, contact days may be asymmetric in a dyad, that is, the contact days for $m_1 \rightarrow m_2$ may be different from the contact days for $m_2 \rightarrow m_1$. Using percentages enabled me to have the same scale for contact days across time frames of different lengths.

The fourth measure of tie strength is duration of contact. For a given time frame, I calculated the difference between the first day and the last day that member m_1 mentioned member m_2. Then I converted this number to a percentage of the total days for the time frame, to allow me to have the same scale across different time frames. Duration of contact can also be asymmetric in a dyad.

I calculated four measures of tie strength between different types of pairs as dependent variables and mediators for the hypotheses. These pair types include the tie from a buddy to the partner, the tie from a buddy to everyone else in the group except the partner on average, and the tie from a member (buddy's partner) to other people in the group on average.

Analyses Approach

Information about goal attainment (abstinence) from online support group members was gathered at the end of each time frame. Each record in my final data set contains information about member m_1 during a time frame. I only included members for whom the abstinence data exists for that time frame. Each record contains data about m_1's group identification number (id), time frame, m_1's member id, m_1's buddy activity level, tie strength from m_1 to other people in the group on average (four tie strength measures), tie strength from m_1's buddy to m_1 (four tie strength measures), tie strength from m_1's buddy to other people except m_1 in the group on average (four tie-strength measures), and m_1's abstinence status.

To test the first hypothesis (H1), I examined the effect of m_1's buddy activity level on m_1's abstinence status. I used a generalized linear mixed model with a binary logistic regression for the dependent variable (DV) in SPSS. Since more than one record for each member may exist (for different time frames), time frame was specified as a repeated measure in the model. The DV was m_1's abstinence status. The fixed effects were m_1's buddy activity, time frame, and the interaction between them time frame × m_1's buddy activity. The random effect was group id. If there was a significant interaction effect, then I did follow up tests to see how the effect of m_1's buddy activity was different on the DV within each time frame.

To test the second hypothesis (H2), I examined the effects of m_1's buddy activity level in the group on the tie strength from m_1's buddy to m_1, using the four tie strength measures. For this test, I used the generalized linear mixed model with time frame as a repeated measure and the dependent variable being one of the four tie strength measures (frequency per day, reciprocity days, contact days, or duration of contact) for the tie from m_1's buddy to m_1. Similar to the first hypothesis, the fixed effects were m_1's buddy activity, time frame, and the interaction between them time frame × m_1's buddy activity. The random effect was group id. To test this hypothesis, I ran four separate tests in SPSS, one for each tie strength measure.

For H3, I compared the tie that m_1's buddy made with m_1 with the ties that m_1's buddy made with other members in the group except m_1 on average, using the four tie strength measures. These are two levels of pair type for the records related to m_1's buddy that were used for this hypothesis. For this test, I used the generalized linear mixed model with the time frame and pair type as repeated measures and the dependent variable being one of the four tie strength measures (frequency per day, reciprocity days, contact days, or duration of contact). The fixed effects were the pair type with two levels (tie from m_1's buddy to m_1 and tie from m_1's buddy to others except m_1 (mean)), time frame, and the interaction between them (time frame × pair type). The random effect was group id. For each tie strength measure I ran a separate test.

For the fourth hypothesis (H4), I examined the effects of m_1's buddy activity level on the tie strength from m_1 to other people in the group (on average), using the four tie strength measures. The setup for this hypothesis was exactly similar to the setup for the second hypothesis except that for H4 the DV was one of the four tie strength measures from m_1 to all members in the group instead of the tie strength from m_1's buddy to m_1 which is the DV in H2. To test this hypothesis, I ran four separate tests in SPSS, one for each tie strength measure.

To test the fifth hypothesis (H5), I evaluated the mediation role of ties from m_1's buddy to m_1 and from m_1 to other people in the group (which were tested in H2 and H4) in the relation between m_1's buddy activity level and m_1's abstinence status (which was tested in H1). To test this hypothesis, I conducted mediational tests using the Hayes macro for SPSS, Model 85, with 5,000 bootstrap samples (Hayes 2018). I conducted a separate test for each measure of the tie strength. For each test, the DV was m_1's abstinence status, the independent variable was m_1's buddy activity level, mediators were the tie from m_1's buddy to m_1 and the tie from m_1 to other group members (on average), and the moderator was the time frame.

Empirical Results

Descriptive Statistics

The average age for all 720 members in the program was 39.28 (SD=9.5) years. They were mainly female (80%), white (81%), and educated (67% college and above vs 33% high school and below); 59% were in a relationship; and 59% were employed.

The 720 individuals in the 36 online support groups posted 73,935 times to their groups over 90 days (mean=102.69, SD=255.22; median = 39.5); 55% of the posts mentioned at least one other individual in the group, 45% did not; and overall the average number of mentions per post was 0.81 (SD=1.17). Out of the 720 individuals, goal attainment (abstinence from smoking) data was obtained for 606 individuals during time frame 1 (84%) and 678 individuals during time frame 2 (94%). Table 1.2 shows the mean, standard deviation, and correlations for the four tie strength measures, involving all potential ties for the two time frames (N = 27360). 20983 records out of these 27360 potential ties were 0 (77%).

Table 1.2 Statistics for tie strength measures

	Mean	Standard Deviation	1	2	3
1. Frequency per day (mean)	0.07	0.42			
2. Reciprocity days (%)	0.78	2.86	0.71*		
3. Contact days (%)	2.78	8.69	0.75*	0.81*	
4. Duration of contact (%)	7.48	21.42	0.50*	0.62*	0.83*

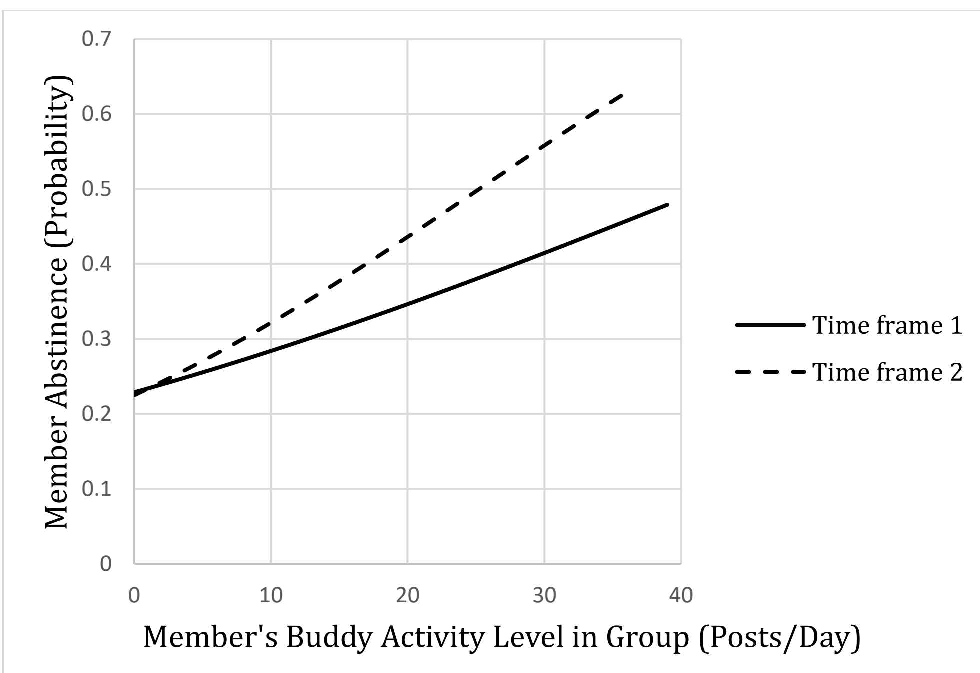
Notes. Mean, standard deviation, and correlations for the tie strength measures (N=27360), *: p<0.001

After aggregating data for each member who had available abstinence data during a time frame, I had a table of 1284 records for two time frames. Each record contains data for one member in one time frame. Records contain the information about the member's buddy activity level in the group during the time frame, the member's abstinence status at the end of the time frame, the values of tie strength measures from the member to other people in the group on average (used for test of H4 and H5), from the member's buddy to the member (used for test of H2, H3, and H5), and from the member's buddy to non-buddies on average (used for H3). For H3, I transformed the 1284 records from the wide format to the long format (based on the pair type with 2 levels). The total number of records in the long format for H3 was 2568 (1284×2).

Test of H1

H1 predicts that those members whose buddies are more active in the group are more likely to achieve their goal (abstinence in this context). The result of a generalized linear mixed model with binary logistic regression target suggested that having a buddy who is more active in the group is related positively to the goal attainment ($F(1, 1280)=4.76$, $p=0.029$). The interaction term for time frame × buddy activity was not significant ($F(1,1280)=0.31$, $p=0.575$). The main effect for time frame was also not significant ($F(1,1280)=0.02$, $p=0.877$). This result supports hypothesis 1. See Figure 1.2. Results are shown for time frames to simplify the comparison with other figures below.

Figure 1.2 Member's goal attainment



Test of H2

H2 predicts that the tie strength from the member's buddy to the member is stronger for those members whose buddies are more active in the group. The results of linear mixed model tests support H2 for all tie strength measures. For each measure that the interaction term time frame × member's buddy activity was significant, I did follow up tests to evaluate the effect of member's buddy activity on the tie strength measure within each time frame.

The result of the test for the frequency per day measure shows that the interaction term for time frame × member's buddy activity was significant ($F(1,1280)= 337.98$, $p<0.001$). For time frame 1, a member's buddy who was more active in the group mentioned the member significantly more than a buddy who was less active in the group ($t=20.89$, $p<0.001$). For time frame 2, the effect was weaker but still significant ($t=12.98$, $p<0.001$).

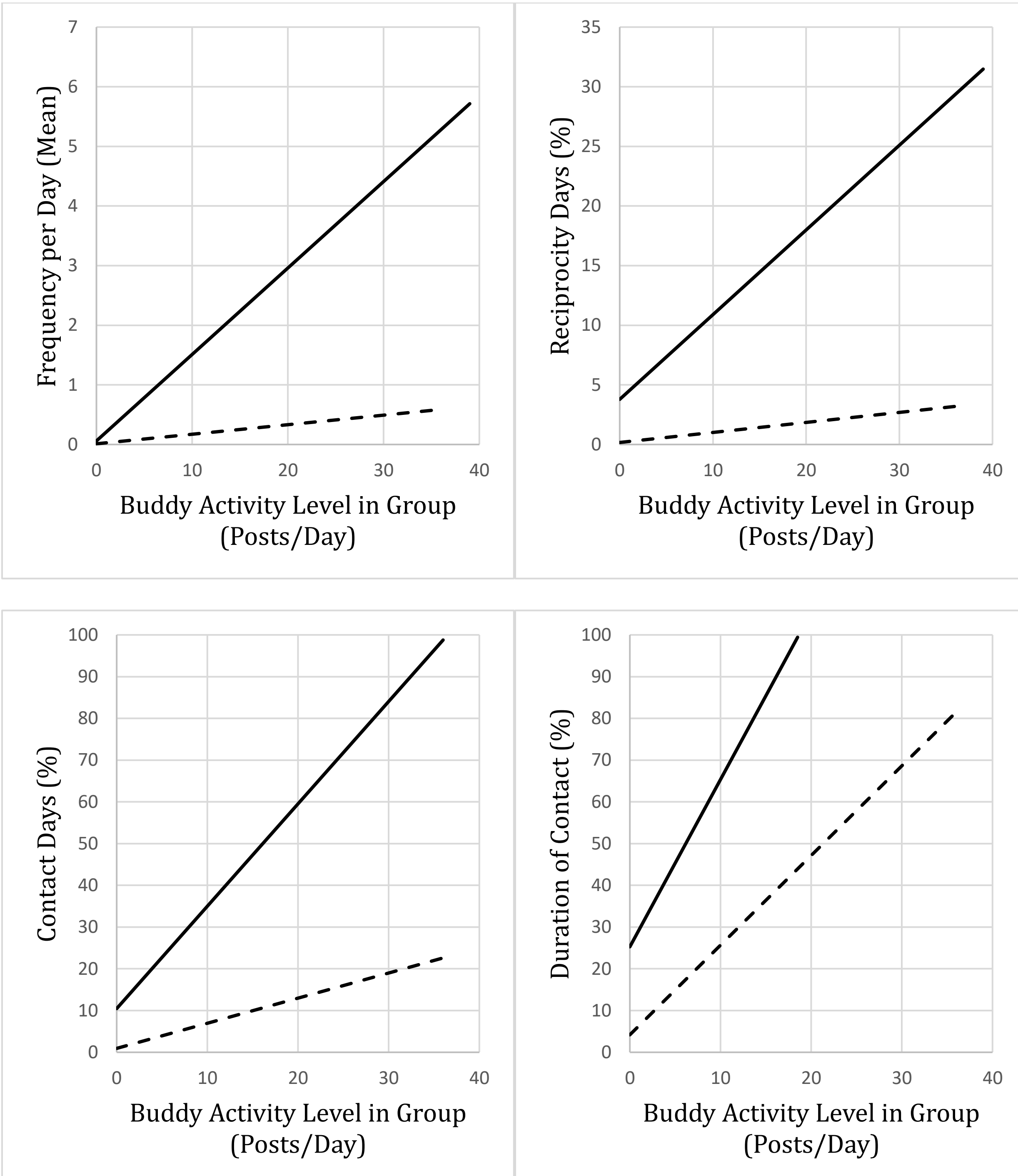
For the reciprocity days measure, the test result shows that the interaction term for time frame × member's buddy activity was significant ($F(1,1280)=75.08$, $p<0.001$). The follow up test for time frame 1 shows that the longest consecutive days in which a member's buddy had reciprocated relations with the member was significantly greater if the member's buddy was more active in the group, as compared to a buddy who was less active in the group ($t=9.96$, $p<0.001$). The effect was weaker for time frame 2 but it was still significant ($t=7.17$, $p<0.001$).

The result of the test for the contact days measure shows that the interaction term time frame × member's buddy activity was significant ($F(1,1280)=98.03$, $p<0.001$). For time frame 1, a member's buddy contact days with the member amounted to a significantly greater percentage of the time frame days if the member's buddy was more active in the group, as compared to a buddy who was less active in the group ($t=14.04$, $p<0.001$). For time frame 2, the effect was weaker but still significant ($t=9.50$, $p<0.001$).

For the duration of contact measure, the result of the test shows that the interaction term time frame × member's buddy activity was significant ($F(1,1280)=16.23$, $p<0.001$). For time frame 1, the duration of the relationship from a member's buddy to the member was a significantly greater proportion of the time frame days for the member whose buddy was more active as compared to being less active in the group ($t=11.41$, $p<0.001$). For time frame 2, the effect was weaker but still significant ($t=9.03$, $p<0.001$).

Overall, the four tie strength measures from the member's buddy to the member were significantly stronger for a member whose buddy was more active in the group. See Figure 1.3.

Figure 1.3 Tie from member's buddy to the member



Notes. Solid line = time frame 1, dotted line = time frame 2

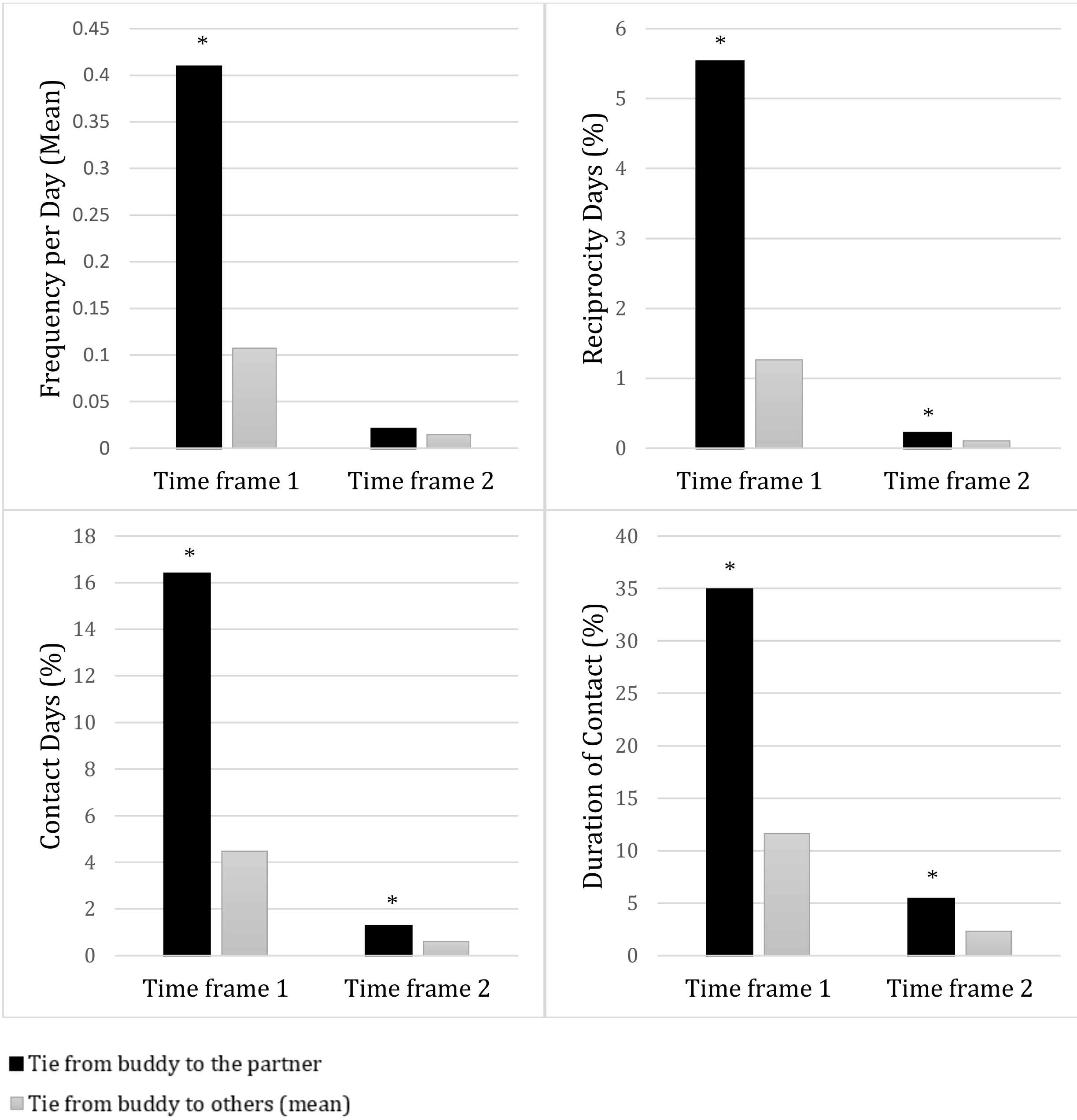
Test of H3

Analysis for H3 confirms that buddies formed especially stronger ties with their partners as compared to other people (non-buddies) in the group, especially during time frame 1. The interaction term was significant for all four tie strength measures frequency per day ($F(1,2564)=57.04$, $p<0.001$), reciprocity days ($F(1,2564)=170.45$, $p<0.001$), contact days ($F(1,2564)=165.75$, $p<0.001$), and duration of contact ($F(1,2564)=121.22$, $p<0.001$). I did follow up tests to compare the ties within each time frame.

In time frame 1, there was a significant difference between the tie from the buddy to the partner and the tie from the buddy to other people in the group (mean) in terms of frequency of mentions per day ($t=7.77$, $p<0.001$), reciprocity days ($t=13.49$, $p<0.001$), contact days ($t=14.01$, $p<0.001$), and duration of contact ($t=13.85$, $p<0.001$).

In time frame 2, the difference between the tie from the buddy to the partner and the tie from the buddy to other people in the group (mean) for frequency per day measure was not significant ($t=1.37$, $p=0.171$). The difference between the tie from the buddy to the partner and the tie from the buddy to other people in the group was weaker than time frame 1 but still significant for reciprocity days ($t=3.13$, $p=0.002$), contact days ($t=3.27$, $p=0.001$), and duration of contact ($t=4.18$, $p<0.001$). See Figure 1.4.

Figure 1.4 Tie to the buddy versus tie to a non-buddy (mean)



Notes. Comparison between the ties that buddies form with their partners and with non-buddies on average, *: p < 0.001.

Test of H4

H4 predicts that the tie strength to other people in the group (mean) is stronger for those whose buddies are more active in the group as compared to those whose buddies are less active in the group. The results of four linear mixed model tests support this hypothesis, especially for time frame 1. For each measure that the interaction term for time frame × member's buddy activity was significant, I did follow up tests to evaluate the effect of member's buddy activity on the tie strength measure within each time frame.

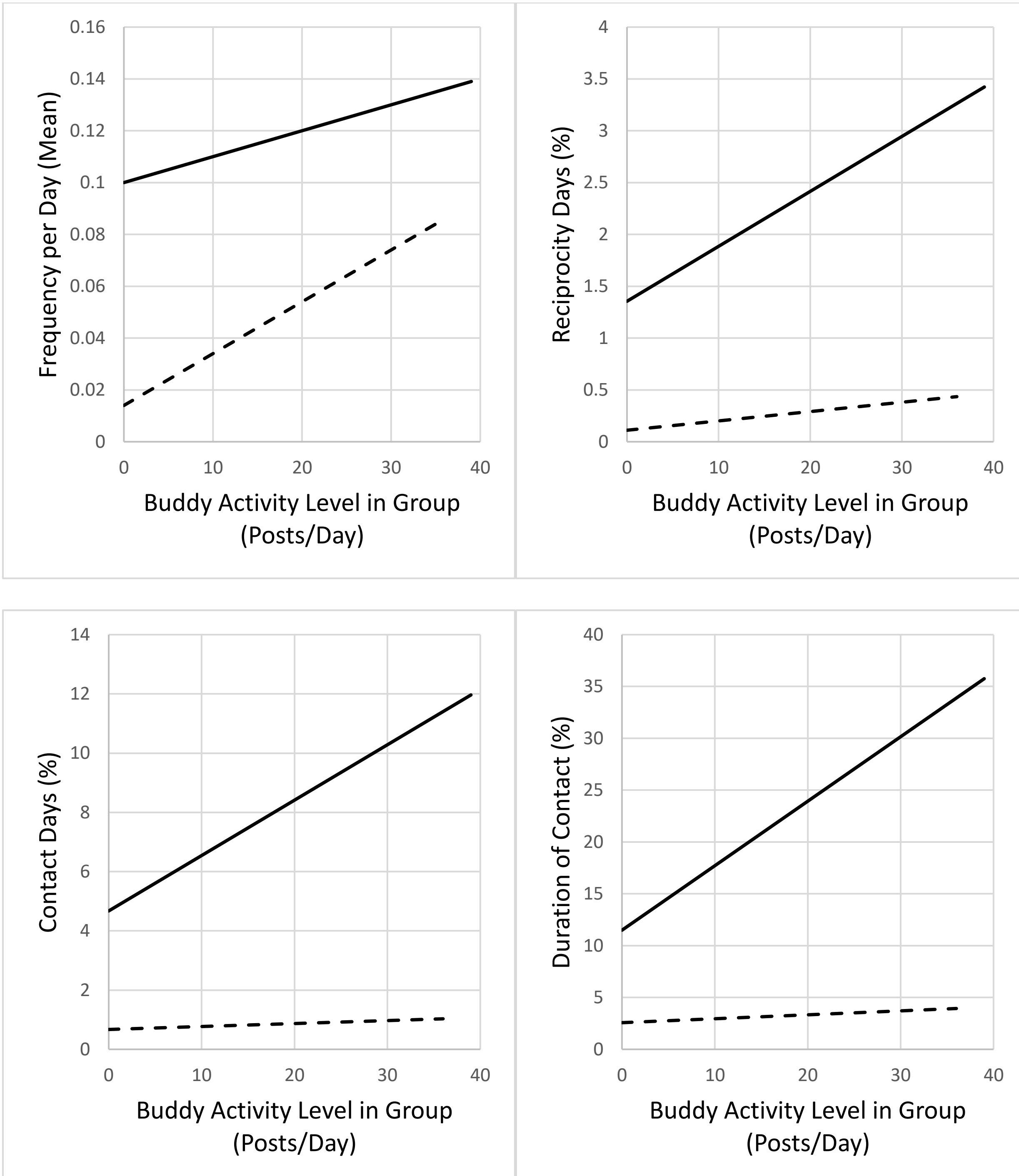
The result of the test for the frequency per day measure shows that the interaction term for time frame × member's buddy activity was significant ($F(1,1280)=7.88$, $p<0.001$). For time frame 1, a group member whose buddy was more active in the group mentioned other people in the group (mentions mean) significantly more than a member whose buddy was less active in the group ($t=2.47$, $p=0.014$). For time frame 2, the effect was only marginally significant ($t=1.72$, $p=0.087$).

For the reciprocity days measure, the test result shows that the interaction term for time frame × member's buddy activity was significant ($F(1,1280)=13.83$, $p<0.001$). The follow up test for time frame 1 shows that the mean of the longest consecutive days in which a member had reciprocated relations with other people in the group was significantly greater if the member had a buddy who was more active in the group, as compared to a member who had a buddy who was less active in the group ($t=3.00$, $p=0.003$). The effect was weaker for time frame 2 but it was still significant ($t=2.30$, $p=0.022$).

The result of the test for the contact days measure shows that the interaction term for time frame × member's buddy activity was significant ($F(1,1280)=10.10$, $p=0.002$). For time frame 1, a member's contact days with other people in the group (mean) amounted to a significantly greater percentage of the time frame days if the member had a buddy who was more active in the group, as compared to a member who had a buddy who was less active in the group ($t=2.59$, $p=0.010$). For time frame 2, the effect was not significant ($t=0.31$, $p=0.758$).

For the duration of contact measure, the result of the test shows that the interaction term time frame × member's buddy activity was significant ($F(1,1280)=12.61$, $p<0.001$). For time frame 1, the duration of the relationship with other people in the group (mean) was a significantly greater proportion of the time frame days for a member with a buddy who was more active in the group, as compared to a member with a buddy who was less active in the group ($t=3.74$, $p<0.001$). For time frame 2, the effect was not significant ($t=0.38$, $p=0.702$). See Figure 1.5.

Figure 1.5 Tie from a member to other people in the group



Notes. Solid line = time frame 1; dotted line = time frame 2.

Test of H5 (Mediation Test)

H5 predicts that having a buddy who is more active versus less active in the group, by strengthening the tie that the buddy forms with the member and consequently by strengthening the tie that the member forms with all people in the group (mean), would increase goal attainment (chance of abstinence). For each tie strength measure I ran a separate mediation test with time frame as a moderating factor. The results of these tests show that for all four tie strength measures, the tie from the buddy to the partner and as its result the tie from the partner to other people in the group mediated the positive effect of having a buddy who is more active in the group on goal attainment. For all four measures of tie strength and for both time frames, the indirect effect of having a buddy who is more active versus less active in the group on the goal attainment, with both mediators in the model, was significant. Time frame moderated this indirect effect for three tie strength measures in the way that the indirect effect was stronger during time frame 1 compared to time frame 2.

For the frequency per day measure, the indirect effect was significant for both time frame 1 (indirect effect B=0.0368, 95% CI= 0.0112, 0.1024) and time frame 2 (indirect effect B=0.0040, 95% CI= 0.0008, 0.0138). The indirect effect was weaker in time frame 2, resulting in negative index of moderated mediation. The confidence interval doesn't include 0, meaning that this negative index is significant (B=-0.0328, 95% CI= -0.0928, -0.0092).

For reciprocity days, the indirect effect was also significant for both time frame 1 (indirect effect B=0.0272, 95% CI= 0.0136, 0.0543) and time frame 2 (indirect effect B=0.0032, 95% CI= 0.0004, 0.0089). The indirect effect was weaker in time frame 2, resulting in significant index of moderated mediation (B=-0.0240, 95% CI= -0.0497, -0.0105).

For contact days, the indirect effect was significant for time frame 1 (indirect effect B=0.0285, 95% CI= 0.0136, 0.0571) and time frame 2 (indirect effect B=0.0069, 95% CI= 0.0028, 0.0164). The indirect effect was weaker for time frame 2, resulting in significant index of moderated mediation (B=-0.0216, 95% CI= -0.0474, -0.0071).

For duration of contact, the indirect effect was significant for time frame 1 (indirect effect B=0.0219, 95% CI= 0.0124, 0.0399) and time frame 2 (indirect effect B=0.0118, 95% CI= 0.0057, 0.0248). The indirect effect was consistent across both time frames, resulting in non-significant index of moderated mediation (B=-0.0100, 95% CI= -0.0262, 0.0028). Table 1.3 summarizes the findings for all hypotheses.

Table 1.3 Summary of findings for hypothesis tests

Hypothesis	Result	Figure
H1. More active vs less active buddy in the group → partner goal attainment	Supported	Figure 1.2
H2. More active vs less active buddy in the group → tie strength of buddy with partner	Supported for all four measures of tie strength	Figure 1.3
H3. Tie strength with partner > tie strength with others	Supported for all four measures of tie strength	Figure 1.4
H4. More active vs less active buddy in the group → tie strength of partner with group members	Supported for all four measures of tie strength	Figure 1.5
H5. More active vs less active buddy in the group → tie strength of buddy with partner → tie strength of partner with group → partner goal attainment	Supported for all four measures of tie strength	Figure 1.1

Discussion

Summary and Conclusion

Online support groups have many applications in the fields of health behavior change and collaborative learning. An important characteristic of these groups is that members seek peer social support to help them pursue their goals. Hence, a main requirement for success in these online groups is to have active members who engage with others and provide social support to each other. In this chapter, I studied the effects of a buddy system on goal attainment and on social support in online support groups. Buddies work better if they extend people's social support networks. I tested to see if and how a buddy system that seeks to extend the social support provided in online support groups relates to member engagement with others and goal attainment in those groups.

To evaluate the effects of assigning buddies in online support groups, unique research questions were answered in this chapter. First, I looked at the effect of the buddy system on goal attainment as the main outcome in the online support groups. Then I looked at the dynamics of this effect in online support groups. For this, I evaluated the tie strength from the buddy to the member and from the member to all people in the group for members whose buddies are more active versus less active in the group. To evaluate the overall causal model, I tested the role of the tie from the buddy to the member and the tie from the member to all people in the group (when the former affects the latter) as mediators in the relation between having a buddy who is more active versus less active in the group and member's goal attainment. I also evaluated the role of time frame as a moderator in the mediation tests. In addition, I compared the ties that members form with their partners with the ties that these members form with other people in the group. The tie strength measures that I used to evaluate the strength of online peer ties are quite general. I used members' observable posts in online groups to calculate the tie strength measures, and specifically their mentions of other members.

Results indicated that a buddy system can increase the chance of goal attainment, improve the level of support received by members and consequently improve the level of engagement with others in online support groups for those whose buddies are more active in the group. Those who had more active buddies as compared to those who had less active buddies in the group had higher chance of goal attainment, had stronger ties from their buddies, and had stronger ties to all group members (mean). My analyses confirmed the mediation role of the tie from the buddy to the partner and from the partner to people in the group (while the former affects the latter) in the relation between having a buddy who is more active versus less active and the goal attainment. Also, time frame moderates this indirect effect for three measures of tie strength. As a result, the casual model suggests that members who were more active in the group provided more support to their partners and this made partners more engaged in the relationship with all the people in the group which consequently resulted in a higher chance of abstinence for those partners. My analyses also confirmed the positive effect of the buddy systems in online support groups by showing that buddies who are more active versus less active in the group help their partners more than other people in the group. These findings are robust across all four tie strength measures.

Limitations

The online support groups that I considered in this chapter are goal-based online support groups in which all members share the same goal. Health behavior change groups and collaborative learning groups are some examples of such groups. In these groups, members benefit from the social support provided by other members. Results presented here regarding the benefits of buddies may not be valid for other types of online communities such as online gaming communities or online communities that serve as direct sales or marketing channels for companies.

Also, in this study, a specific method was used to assign buddies in the online support groups. The buddies were assigned based on similarity in age, gender, geographic location, and years of education. Results will likely differ if buddies are chosen in other ways, e.g., through random assignment or chronological order of entrance. Therefore, demographic and location data should be collected a priori to allow buddies to be assigned in the same way as in this study.

CHAPTER 2: Adding a Chatbot to Online Support Groups: The Natural Language Understanding Component

Different definitions can be suggested for software bots based on their application, intention, interfaces, and hosting environments (Lebeuf, Storey and Zagalsky 2017). In a general form, I define software bots as software programs that interact with humans to make a conversation with them and/or to do a service for them using text and/or voice user interfaces (UIs).

The history of software bots goes back to the 1950s when Alan Turing brought up the idea of his test which requires that a human cannot distinguish between a machine and a person who have a conversation over a text channel (Turing 1950). Early software bots were designed around the idea of passing the Turing test. ELIZA was one of the earliest ones which used keyword matching to analyze the input text and used rule-based decisions to generate the output text (Weizenbaum 1966). Another example is ALICE which was initially released in 1995 and was three-time winner of the controversial Loebner Prize which is annually awarded to the most human-like computer program (Wallace 2009; Powers 1998). Mainly after 1990, attempts were begun to enable the software bots to make vocal conversations with humans (Glass et al. 1995; Hemphill, Godfrey and Doddington 1990; Levin et al. 2000; Seneff et al. 1998). Instead of passing the Turing test, the focus in these systems was mainly on language understanding and doing specific services such as providing transportation information for users (Clementino and Fissore 1993; Blomberg et al. 1993; Peckham 1991). However, these systems never got popular like what we know as software bots today.

Advances in AI and the ubiquity of smartphones initiated a new era of software bots in the 2010s (Dale 2016). Software bots that emerged in this era are divided into three main categories. The first group of these bots are called messaging chatbots or simply chatbots. Chatbots are accounts inside instant messaging applications that are controlled by software. People can interact with them individually or they can be added into groups which are created inside messaging applications. Chatbots are usually task-based and provide specific service or information in a limited domain to users (Klopfenstein et al. 2017). For example, from inside your Facebook messenger application, you can interact with the Kayak chatbot (Booking Holdings Inc. 2020) to search for travel related information, or you can use the NBA chatbot (NBA Media Ventures LLC 2020) to get information about your favorite NBA team.

There have been thousands of chatbots developed during the last decade and their popularity is still rising. The main driving factors of this massive growth are popularity of instant messaging applications and simplicity of integrating chatbots inside them. Since 2014, many instant messaging applications such as Messenger, Telegram, Skype, Slack, and Kik have provided infrastructures for chatbot integration (Klopfenstein et al. 2017). In 2017, the number of monthly active users only in Facebook Messenger passed 1.3 billion (Clement 2019). Businesses, organizations, and any online service provider can take the advantage of this potential user-base and they can directly connect to their users or customers inside these instant messaging applications instead of developing their own applications or attracting customers to their websites. Studies have shown the usefulness of using chatbots for different tasks such as controlling Internet of Things (IoT) devices (Mardini, Khamayseh and Smadi 2017), tracking food for health purposes (Graf et al. 2015), and enabling parents to get information about their children's presence at school (Chaniago and Junaidi 2019). For more complex services such as a teaching assistant chatbot, chatbots are still in their early stages (Smutny and Schreiberova 2020).

The interaction with chatbots is usually provided through text and other user interface (UI) components such as buttons that are provided by the hosting environments. Since the main purpose of chatbots is to provide services in a limited domain, they do not necessarily need to have natural language processing (NLP) components if it is not required for their service. There are different online platforms that provide the necessary infrastructures for developers to build chatbots. Pandorabots, one of these platforms, claims that more than 275K developers have used this platform to create more than 325K chatbots (Pandorabots Inc. 2020).

Intelligent virtual assistants (Dale 2016) are the second group of software bots. They are usually standalone applications with voice and text interfaces that can do several tasks such as setting a reminder, searching on the web, and playing music (Hoy 2018). Apple's Siri is the first popular intelligent virtual assistant which was introduced for iPhones in 2011 (Gross 2011). Other important software bots in this category are Microsoft's Cortana, Google Assistant, and Amazon's Alexa. Unlike messaging chatbots, the domain of tasks that these virtual assistants can do is not limited. The big companies behind these virtual assistants are constantly adding to the bot capabilities and also to their domain of activity. Moreover, they have provided frameworks which enable third parties to develop new skills for the bots (Amazon.com Inc. 2020; Wang 2019; Google LLC 2020). These skills can be used by all users of the intelligent virtual assistants.

Although intelligent virtual assistants are voice driven, and users can hold conversations with them, their main focus is on doing tasks. There are some practices to develop software bots called social chatbots that can make emotional connection with users (Shum, He and Li 2018). The social chatbots can usually do some tasks, but their main goal is to be a virtual companion for users (Shum, He and Li 2018). Replica and Microsoft's XiaoIce (Zhou et al. 2020) are two important examples of the software bots in this category. Social chatbots can be standalone applications like Replica, or they can be software-controlled accounts in instant messaging applications like XiaoIce. Table 2.1 summarizes similarities and differences between different categories of software bots.

Table 2.1 Comparison between different categories of software bots

	Chatbots	Intelligent virtual assistants	Social chatbots
Focus	Doing specific tasks	Doing several tasks, can learn new skills	Emotional and continued conversations
Implementation platform	Instant messaging applications	Standalone applications	Standalone applications/ Instant Messaging applications
Domain	Limited	Not limited	Not limited
User interface (UI)	Mainly text and UI components in the IM application	Voice, text, image	Voice, text, image
Example	Kayak, NBA, Sephora, and UNICEF USA in Facebook messenger	Apple (Siri), Google (Google Assistant) Microsoft (Cortana), Amazon (Alexa)	Replika, Microsoft's XiaoIce

The interaction with all types of the above software bots is initiated by users. There is another type of software bots usually known as social bots that try to emulate the behavior of humans and possibly affect their behavior by automatically generating content and interacting with them on social media (Ferrara et al. 2016). Although these social bots have many similarities with the above three categories from a technological point of view, their goal is completely different. These bots can be considered as the fourth category of software bots with negative goals. Social bots usually use the popularity of social media for negative purposes such as distributing low-credibility information (Shao et al. 2018) for political purposes (Stella, Ferrara and De Domenico 2018; Hegelich and Janetzko 2016) and affecting health policies (Broniatowski et al. 2018). Most studies about social bots are related to the adversarial role of these bots in social networks (Boshmaf et al. 2013) and also to methods for identifying these bots (Davis et al. 2016; Yang et al. 2019).

In this chapter, I present the design of a software bot as a new component to online support groups to improve their performance for members. This bot has a limited domain of activity, has natural language processing capabilities, and can be implemented in an instant messaging application. Therefore, I put it under the first category of software bots, which is chatbots. The main motivational factor for using chatbots is the convenience of getting assistance or information (Brandtzaeg and Følstad 2017). For some people, a bot also has a social value. They see the bot as an interaction mechanism that can be used to improve their social interactions with other people (Brandtzaeg and Følstad 2017). These uses make a chatbot a strong addition to online support groups where members seek informational and motivational support from others to achieve their goal.

Context

Setting

My chatbot is designed to be implemented into online support groups with a specific goal of behavior change. The chatbot was trained and tested on the data from the

Tweet2Quit project. I have data from 14 online support groups of 20 members each. These groups are a subset of the same groups that were studied in Chapter 1. Members in these groups were smokers who joined the program to quit smoking. The groups were created in Twitter. In each group, all members were added to the group at the same time and were asked to support each other for 90 days by tweeting to the group.

Roles of the chatbot

The chatbot that I introduce in this chapter has informational and motivational roles in online support groups. All members in the group know that messages come from a machine not a human. I want the chatbot to behave like an active member who helps other members in the group, especially when it is related to the goal that the group is designed for. I also want the chatbot to encourage others in the group to post messages.

Situations in which I want the chatbot to play the informational role are when members in the group ask questions or raise issues related to practices defined by group designers and that are also in the direction of the defined goal for the group. For example, in the context of smoking cessation support groups, this can be a question or an issue about nicotine replacement therapy (NRT). A previous study of online smoking cessation support groups shows that specific message contents such as nicotine patches are related to the goal attainment of members (Pechmann et al. 2015). If a member asks such a question or raises an issue, I want the chatbot to react by sending related information which could help the member to know what to do or how to handle the issue.

The chatbot plays its motivational role when one member provides some information about his/her accomplishment or temporarily failure. The chatbot is expected to praise accomplishments to make them stronger and more consistent in the member's behavior and to provide an incentive for other people in the group to achieve similar accomplishments. For messages about failures, the chatbot is expected to encourage members to try again and recover from failures.

These are common scenarios in online support groups so that handling them not only creates discussions in groups and keeps members more engaged, but also directly helps members to more easily achieve their goals for which they joined the group.

Design

Chatbot Design

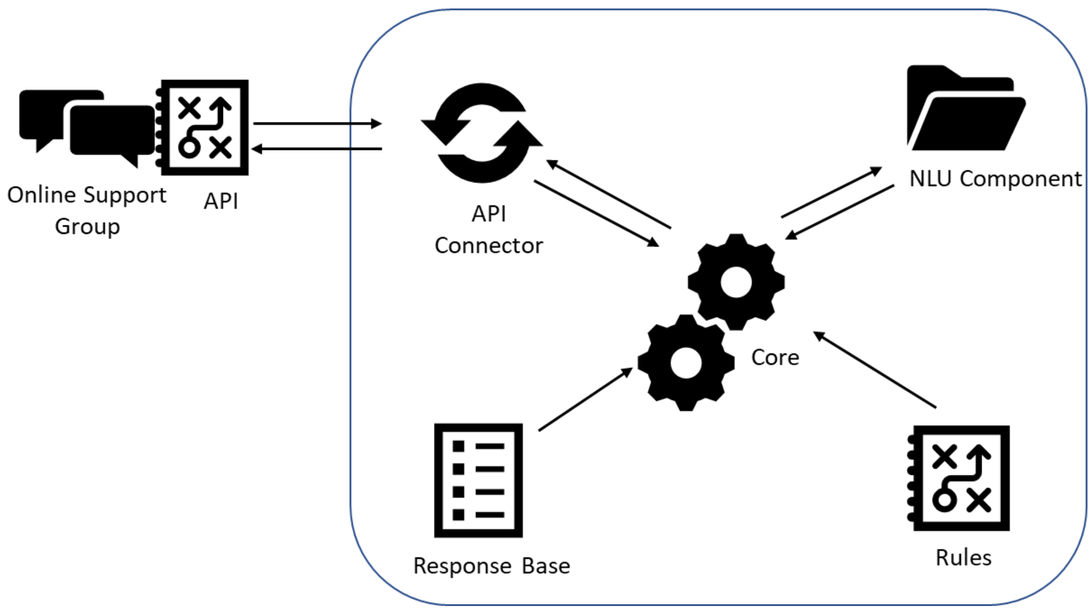
The main task of the chatbot presented here is to provide information or motivation for members in online support groups by responding to their messages in the group. Figure 2.1 shows the high-level software component design of this chatbot.

The instant messaging application which hosts the online support groups (for example Twitter) provides an API (Application Programming Interface) which allows a computer program to connect to the group and read messages from the group or send new messages to the group. The chatbot connects to the API through its API connector component.

The Natural Language Understanding (NLU) component is the content analysis component that receives a message from the core component and uses machine learning models to predict the intent type of the message. This component needs to be trained initially so that it can be used for predicting intent types.

The chatbot needs some rules which specify when and how it should respond to an incoming message based on the intent it predicts. The rules are defined in the Rules component. All responses that the chatbot selects from based on the defined rules to respond to incoming messages are stored in the Response Base. The Core component is responsible for all interactions between other components and the dynamic flow of the information in the chatbot. It implements the rules and handles the logical control of the chatbot.

Figure 2.1 Chatbot components



Based on the chatbot's defined roles, it is expected to instantly respond to some messages which are posted to the group by members. The dynamic flow of information is as follows. The chatbot monitors the group and receives a message which is sent to the group by a human group member. Then it uses the NLU component to recognize the main intent of the message. After that, it uses the defined rules to decide whether to respond to the message or not. If it decides to respond to the message, it retrieves an appropriate response from the Response base and sends that response to the group as a reply to the initial message.

The performance of the chatbot highly depends on the performance of the NLU component for predicting correct intents of messages. If the chatbot fails to understand the main intents of messages, then it cannot respond accordingly. The remainder of this chapter is mainly about the NLU component for predicting intents of new incoming messages. The design and development of other components are straightforward and are not covered here.

Intents

Intent is the core meaning and the main purpose of an expression. Intents are the main information in the chatbot system which are used to make decisions about how to respond to the incoming messages from members. They are identified by the NLU component and are used in the Rules component. For Tweet2Quit smoking cessation online support groups, 26 intents are considered. These intents are categorized into the following main groups:

Using Nicotine Replacement Therapy (NRT) correctly. These are important intents related to the side effects of using NRT products including patches, gums, and lozenges, especially when they are not used correctly. I want the chatbot to identify messages with these intents and respond to them by providing information on how to use the NRT products correctly. Table 2.2 shows seven intents of this category with some examples and the appropriate type of chatbot response to them.

Table 2.2 Intents about using NRT correctly

Intent	Description	Examples	Bot response
NRT_howToUse	Asks question or gives instructions about how to use NRT products	How often should I use the lozenges? How do I put on the patch? Wear them on different places. Chew it for a little bit and put between cheek and gum.	Information about using NRT products
NRT_stickIssue	States NRT patch won't stick	I was scared it would come off. My patch won't stay on.	Information about making the patch stick
NRT_dreams	States sleep issues related to NRT patches	Did you have a vivid dream too? I've heard that it isn't a good idea to sleep with them on. Something about dreams and restless sleep.	Information about not using the patch while sleeping
NRT_skinIrritation	States NRT patch causes skin irritation	When you put on patch, it itches really bad. Itchy for an hour.	Information about how to prevent NRT patch skin irritation
NRT_mouthIrritation	States NRT gum/lozenge has bad taste, irritates throat, or causes sense of burning or spicy	The gum has a strong nicotine taste. The lozenges do not taste good. The gum burns my mouth. It hurts my throat.	Information about preventing gum or lozenge mouth irritation
NRT_nauseous	States NRT makes nauseous or gag	It made me nauseous. I want to throw up.	Information about preventing nausea from NRT products
NRT_OD	States NRT is too strong and causes overdose	The 21mg patches are making me sweat and feel bad. It was giving me too much nicotine. I am doing better with the 14 MG.	Information about correct dose of NRT to use

Efficacy of NRT. These intents are mainly related to the efficacy of NRT products. The response of the chatbot to these types of intents may be informational or motivational. Table 2.3 shows three intents of this category.

Table 2.3 Intents about efficacy of NRT

Intent	Description	Examples	Bot response
NRT_itWorks	States NRT works	The patches work if you're determined. Lozenges are good. I still use the gum for cravings. The patch every day and a piece of gum when I get a severe craving.	Provide motivation by praising the continuous use of NRT
NRT_don'tWork	States NRT doesn't work, don't like it, or no longer need it	The patches aren't working that great for me. The NRTs aren't for my quitting. The gum has not made any difference at all. I don't like the gum. I don't need patches or nicotine gum.	Information about NRT efficacy and encourage to continue to use NRT
NRT_runOut	States NRT will, or did, run out	I will be running out of patches sooner than planned. I am almost out of patches.	Information about how to get more NRT

Negative emotions. These intents are mainly related to the negative emotions such as cravings, stress, and tiredness that arise during quitting smoking. For the messages with these intents, I want the chatbot to respond by providing information and motivation that help members with managing these negative feelings. Table 2.4 shows five intents of this category.

Table 2.4 Intents about negative emotions

Intent	Description	Examples	Bot response
fail	States smoking cigarette	False start this morning. I had a miserable day! Had 2 today.	Provide motivation to try again, don't give up
cravings	States feels craving to smoke	Cannot stand the cravings. Want to smoke Wanting a cigarette.	Information about how to overcome cravings
stress	States feeling stressed	Need something for stress. Lots of stress factors.	Information about how to manage stress

scared	States feeling nervous about NRT or quit	I am very nervous to use the patch. I am scared to fail. I am so scared.	Information about how to overcome being scared
tiredness	States feeling tired	I feel like I am about to fall asleep. I am out of energy.	Information about how to overcome tiredness

Positive actions. These intents are about positive actions or accomplishments related to the goal for which members joined the group. For these intents, I want the chatbot to respond by praising and motivating members to continue doing positive things. Table 2.5 shows five intents of this category.

Table 2.5 Intents about positive actions

Intent	Description	Examples	Bot response
support	Supporting, praising, motivating other members	Never give up. You can do this! Good job. Congrats on not smoking.	Provide motivation by supporting more
smokeFree	States success in being smoke free	Still smoke free. It has been 13 days since my last smoke. I've had some severe cravings today but worked through them. Still going strong not smoking.	Provide motivation by congratulating for being smoke free
quitDate	States something about quit date	Monday October 8th is my quit date. My quit date is getting closer.	Provide motivation by wishing good luck on the quit date
costs	States the cost benefits of quitting	Money is another huge reason I want to quit. Saving for grandsons.	Provide motivation by stressing the benefits of saving money
smokingLess	States success in smoking less	A pack a day to 4 cigs. Max of three cigs a day.	Provide motivation by congratulating for smoking less and encourage to quit completely

Health. These intents are about different health topics such as weight gain. For these intents, I want the chatbot to provide related helpful information and tips that are useful for members. Table 2.6 shows four intents of this category.

Table 2.6 Non-triggering intents

Intent	Description	Examples	Bot response
health	State something about the relation between smoking and health	My health keeps me motivated to not smoke. I want to quit to be healthier.	Information about smoking and health
weightGain	State something about gaining weight or eating more	I tend to want to snack more. Picked up eating too much.	Information about exercise and diet
ecigs	State something about e-cigarettes	I did vape. Menthol juice.	Information about e-cigarettes
cigSmell	Complains about cigarette smell	It smelt so disgusting. I don't like the smell of cig.	Information about how to get rid of cigarette smell

Greetings and others. In addition to the above intents, I want the chatbot to identify all messages with the intent of greetings. For this intent, I want the chatbot to respond by greeting back. Except these 25 mentioned intents, all other potential intents are categorized as one non-triggering intent named as "others". I don't want the chatbot to do anything when it receives a message with an intent identified as "others".

Natural Language Understanding (NLU) Component

In this part, I describe the structure of the NLU component for the Tweet2Quit smoking cessation online support groups. However, the ideas and the model are extendable to other types of online support groups. The NLU component in my chatbot is basically a multi-class classifier for intents. I use machine learning (ML) models to predict intents of messages that are posted by members to online support groups.

I needed labeled data to train my intent classifier. I had 14 groups with labeled data. 75% of the data from these groups was randomly selected for the purpose of training. Since the data was imbalanced across different intents, I used the SMOTE (Chawla et al. 2002) method for oversampling the training data to improve the performance of the multi-class classifier. SMOTE is a widely used oversampling approach which creates synthetic minority class samples to make the dataset more balanced.

Before training the model for intent prediction, I needed to clean and prepare the data. I performed several operations on the text of messages to prepare them better for ML models. This includes removing mentions and links, removing unnecessary characters such as extra white spaces, separating emojis (both Unicode emojis and text emojis) from the main text, making the text lower case, and removing stop words such as “the”, “a”, “an”, “to”, “and”, and “as”. I also considered Snowball stemming (Porter 2001) of messages which algorithmically removes some suffixes such as “ive” and “ly” from words and WordNet (Princeton University 2010) lemmatization of messages which replaces different inflected forms of a word with one word. But, Snowball stemming and lemmatization did not improve the performance of the model and were removed from the final version of the code.

After cleaning the text data, I used the tf-idf (term frequency-inverse document frequency) vectorizer to transform the text of messages into numbers which can be used as independent variables in the learning model. This is a popular method with good performance that considers the frequency of each term in a message and also the reverse frequency of the term among all messages. I considered terms consist of up to three words for the vectorizer. Using this method, I had 9168 features (independent variables) for each record.

I used the random forest model for predicting the intents of messages. To test the performance of the intent classifier, I used a randomly selected 25% of the labeled data from the 14 groups. The complete Python code of my intent classifier is presented in the Appendix.

Result

Descriptive statistics

I had 16136 labeled messages from fourteen groups. 25% of these 16136 messages were stratified randomly selected for testing and the remaining were used for training. Overall, I had 4034 records for testing and 12102 records for training before oversampling. Table 2.7 shows the frequency of each intent out of 16136 initial records. As is shown in this table, the “others” intent accounts for more than 54% of all messages. On the other hand, some intents such as NRT_runOut and NRT_OD are very rare with frequency of less than 1% of all messages. This makes the data highly imbalanced. For training, I used oversampling to have the same number of messages for each intent except “others”. After oversampling, I had 50615 records for training, 6490 messages for others and 1765 messages for each of 25 other intents. Although oversampling handles the issue of imbalanced data, it does not address the problem raised from having a very low number of records for some intents. I do not expect the model to have a good performance on these very rare intents. These intents are important for the application of the chatbot and I want to keep them in my model. This problem can be solved by adding more data from other groups to the dataset which can result in training the chatbot better.

Table 2.7 Intent frequency of records (N=16136)

Category	Intent	Number of messages	Percentage
Using NRT correctly	NRT_howToUse	109	0.7
	NRT_stickIssue	78	0.5
	NRT_dreams	77	0.5
	NRT_skinIrritation	54	0.3
	NRT_mouthIrritation	52	0.3
	NRT_nauseous	45	0.3
	NRT_OD	25	0.2
Efficacy of NRT	NRT_itWorks	274	1.7
	NRT_don'tWork	91	0.6
	NRT_runOut	16	0.1
Negative emotions	fail	350	2.2
	cravings	276	1.7
	stress	204	1.3
	scared	119	0.7
	tiredness	49	0.3
Positive actions	support	2353	14.6
	smokeFree	1165	7.2
	quitDate	511	3.2
	costs	139	0.9
	smokingLess	88	0.5
Health	health	310	1.9
	weightGain	143	0.9
	ecigs	116	0.7
	cigSmell	111	0.7
Greetings	greetings	728	4.5
Others	others	8653	53.6

Intent classification

From 4034 test records, 2910 records were classified correctly. This gives the overall accuracy of 72%. Out of 4034 records, 1871 records had a triggering intent (all intents except "others"). My intent classifier predicted a triggering intent for 1008 records from which 821 predictions were correct. This provides the precision of 81.4% for triggering intents, meaning that if my classifier predicts a triggering intent for a message, the probability that the prediction is correct is more than 0.81. Also, the recall (sensitivity) for the non-triggering intent ("others") is more than 96%, meaning that less than 4% of messages with the non-triggering intent are predicted incorrectly as having a triggering intent. On the other hand, the recall for the triggering intents is 44%, meaning that my classifier predicts a correct triggering intent for about 44% of messages with a real triggering intent. Figure 2.2 shows the complete confusion matrix for the test data of my intent classifier. Rows show actual intents and columns show predicted intents. The numbers on diagonal show the number of correct predictions for each intent. In the next part, I discuss more why I am interested to have higher precision as compared to higher recall for triggering intents. This results in having many non-zero numbers in the last column of Figure 2.2 which shows many messages with triggering intents are classified as "others" which is the non-triggering intent.

Figure 2.2 Confusion matrix for the intent classifier (N=4034)

Actual \ Predicted	NRT_howToUse	NRT_stickIssue	NRT_dreams	NRT_skinIrritation	NRT_mouthIrritation	NRT_nauseous	NRT_OD	NRT_itWorks	NRT_don'tWork	NRT_runOut	fail	cravings	stress	scared	tiredness	support	smokeFree	quitDate	costs	smokingLess	health	weightGain	ecigs	cigSmell	greetings	others
NRT_howToUse	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	23
NRT_stickIssue	1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	16
NRT_dreams	0	0	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	10
NRT_skinIrritation	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	12
NRT_mouthIrritation	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	11
NRT_nauseous	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9
NRT_OD	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5
NRT_itWorks	0	0	0	0	0	0	0	10	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	58
NRT_don'tWork	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	21
NRT_runOut	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4
fail	0	0	0	0	0	0	0	0	0	0	7	0	3	0	0	3	3	3	0	0	0	0	0	0	0	69
cravings	0	0	0	0	0	0	0	1	0	0	0	17	0	0	0	1	0	0	0	0	0	0	0	0	1	49
stress	0	0	0	0	0	0	0	0	0	0	0	0	25	1	0	0	0	0	0	0	0	0	0	1	0	24
scared	0	0	0	0	0	0	0	0	0	0	0	0	0	11	0	0	0	4	0	0	0	0	0	0	0	15
tiredness	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	10
support	0	0	0	0	0	0	0	1	0	0	0	1	0	1	0	320	8	1	1	0	0	0	0	0	16	239
smokeFree	0	0	0	0	0	0	0	3	2	0	1	7	3	0	0	4	106	2	0	0	1	0	0	0	5	157
quitDate	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	89	1	0	0	0	0	1	1	34
costs	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	18	0	3	0	0	2	0	12
smokingLess	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	19
health	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	3	0	27	0	0	1	0	45
weightGain	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	9	0	0	0	26
ecigs	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	14	0	0	13
cigSmell	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	19	0	9
greetings	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	2	2	0	0	0	1	0	0	128	47
others	0	0	0	0	0	0	0	1	0	0	0	6	1	1	0	18	20	11	2	1	2	2	0	0	9	2089

Table 2.8 shows the aggregated data for true positive (TP), false positive (FP), false negative (FN), support, precision, and recall for each category of intents in the test dataset. For each intent, TP is the number of records that are correctly classified with that intent, FP is the number of records that are incorrectly classified with that intent, FN is the number of records with that intent that are incorrectly classified with a different intent, and support is the number of records with each intent in the test dataset (TP+FN). The aggregated data for TP, FP, FN, and support for each intent category in Table 2.8 are the sums for all intents in

that category. Precision shows what percentage of messages that are predicted as one intent actually had that intent (TP/TP+FP). Recall shows what percentage of messages with one intent are classified correctly with that intent (TP/TP+FN).

Table 2.8 Classification report for the test dataset (N=4034)

Intent category	TP	FP	FN	Support	Precision (%)	Recall (%)
Using NRT correctly	19	1	90	109	95.0	17.4
Efficacy of NRT	11	9	85	96	55.0	11.5
Negative emotions	61	28	189	250	68.5	24.4
Positive actions	533	101	531	1064	84.1	50.1
Health	69	15	101	170	82.1	40.6
Greetings	128	33	54	182	79.5	70.3
Others	2089	937	74	2163	69.0	96.6

For the intent category of using NRT correctly, the classifier had great precision. Although it only caught about 20% of messages with these types of intents, 95% of those messages identified with those intents were classified correctly. For the intent category of efficacy of NRT, the classifier had a moderate performance. It only caught about 10% of messages with these types of intents and only 55% of those which were classified with these intents were classified correctly. The main reason for the weak performance of the classifier for this category is the low number of records in this category which is the lowest among all categories. The classifier had acceptable performance on the negative emotions category with precision of 68% and recall of 24% for this category. The best and the most robust performance of the classifier was for the categories of positive actions, and of health with very high precisions and good recalls. Greetings had also very good precision and recall, but this intent has the least importance compared to other triggering intents. For the non-triggering intent “others”, the recall is very high, and precision is also good which is a great outcome for the classifier. I discuss this more in the next part. Adding more data, especially for intents that have low support, can significantly improve the performance of the classifier.

General Discussion

Findings

In this chapter I presented a design for the chatbot with natural language processing capabilities that can be added to online support groups and provide informational and motivational support for the members to help them more easily achieve their goal for which they joined the group. The chatbot is intended to send a set of discrete one-time messages to the group in response to the members’ messages which seek information or motivation instead of making continuous conversations with members. Unlike social chatbots such as XiaoIce whose goal is mainly to make continuous emotional interactions with users and its success is evaluated by measures such as Conversation-turns Per Session (Zhou et al. 2020), the functionality performance of chatbots is dependent on the quality of the task or the data that they provide for users. The quality of the service provided by the chatbot presented here is highly dependent upon recognizing the intent of members’ messages to which the chatbot wants to respond. This is the responsibility of the NLU component of my chatbot. Therefore, the NLU component determines the performance of the chatbot. This component is basically a multi-class classifier for intents. The rules that define the chatbot behavior depend on the intents that are predicted by the NLU component. In this chapter I mainly focused on the design and development of the NLU component. Developing other components is a straightforward task.

I considered 26 different intents in 7 categories (2 categories have only 1 intent each) for labeling messages in the dataset. While messages with the non-triggering intent account for about 54% of all records, some triggering intents, especially most of the important NRT related ones, have frequency of less than 1% in the studied online support groups. I needed to keep these low-frequent intents because they are important, and I want the chatbot to respond to them. This makes the data from the online support groups highly imbalanced. One method to address the issue of imbalanced data especially when the data set size is not large enough is oversampling. The SMOTE (Chawla et al. 2002) oversampling method worked very well for my case. I used this method to have the same number of records for each intent except "others" for training the NLU classifier.

I intentionally wanted to have more records for the "others" intent in the training dataset because I wanted the intent classifier to give more weight to the non-triggering intent rather than triggering intents which is a good thing for the chatbot. High precision, which shows a low false positive, is much more important than high recall, which shows a low false negative, for the chatbot when I want to identify triggering intents. It is not a big problem if the chatbot misses some triggering intents by predicting them as non-triggering. However, it is very bad to predict a non-triggering intent as a triggering one or to predict a triggering intent as another incorrect triggering intent. For this reason, I made the training data balanced among triggering intents but kept it imbalanced compared to non-triggering intents. The imbalanced pattern of data among triggering and non-triggering records which gives more weight to the non-triggering intent ("others") provides higher precision as compared to higher recall for predicting triggering intents which is in the favor of my chatbot. My intent classifier achieved the total precision of more than 81% for triggering intents on the test data. It can also identify the correct intent for more than 44% of messages with triggering intents. The high precision value for triggering intents shows that most messages with triggering intents which are not classified correctly are classified as the non-triggering intent. This aligns with my goal of higher precision as compared to higher recall for triggering intents. Among all intent categories, the classifier had the best performance for the positive actions and health categories while the most challenging category was the efficacy of NRT. Adding data from new groups to the dataset can improve the classifier performance for all categories especially those with small number of records.

Adding a chatbot is beneficial to an online support group by providing members with additional information and motivation, and also by improving their engagement in the groups. This can help members to achieve their goal more easily.

Limitations and next steps

The intent classifier presented in this chapter is not a general intent classifier. It is trained for online support groups with the goal of quitting smoking. More studies are required to see if this type of intent classifier works fine for other type of online support groups. The classification result is also dependent upon the defined intents. If there is a need to change the set of intents, the labels must be updated, and the result may change.

I used more than 16000 labeled records from 14 online support groups for training and testing. Having access to labeled data for other types of online support groups is necessary for training the chatbot. My data was highly imbalanced. For some important intents such as NRT_runout or NRT_od, I had a very small number of records. This highly affected the performance of the classifier for these rare intents. Data from more groups is required to improve the performance of the chatbot for these intents.

To have an operational chatbot which can be added to online support groups in messaging applications like Twitter, other components of the chatbot must be developed. This includes writing the code for the API connector which enables the chatbot to read the messages that are posted to the group and to send messages to the group, using the intents to define the rules for handling each situation when the bot reads a message from the group, providing appropriate answers to be sent to the group for each situation which are defined by rules, and developing the core component of the chatbot which connects all other components together and implements the business logic of the chatbot.

In this chapter I presented the design for the chatbot and I also developed and tested an intent classifier which enables the chatbot to respond to members' messages appropriately. This chapter provides a framework for future empirical studies evaluating the effects of adding a chatbot to online support groups. Such studies can provide an evaluation of the overall performance of chatbots in online support groups. This includes answering the following questions: What is the effect of adding a chatbot on members' goal attainment in online support groups? What are the effects of adding a chatbot on group dynamics such as interactions between members in the group? How do messages from the chatbot affect the content of humans' messages in the group? Randomized control trial experiments are required to answer to these important questions.

REFERENCES

Amazon.com Inc. (2020), "Alexa Skills Kit." https://developer.amazon.com/en-US/alexa/alexa-skills-kit. Accessed 7 July 2020.

Aral, Sinan and Dylan Walker (2014), "Tie Strength, Embeddedness, and Social Influence: A Large-Scale Networked Experiment." *Management Science*, 60, 6, 1352-1370.

Arguello, Jaime, Brian S. Butler, Elisabeth Joyce, Robert Kraut, Kimberly S. Ling, Carolyn Rosé, and Xiaoqing Wang (2006), "Talk to Me: Foundations for Successful Individual-Group Interactions in Online Communities" in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems,* Rebecca Grinter, Thomas Rodden, Paul Aoki, Ed Cutrell, Robin Jeffries and Gary Olson eds. New York: ACM, 959-968.

Armstrong, Arthur and John Hagel III (1996), "The Real Value of on-Line Communities." *Harvard Business Review*, 74, 3, 134-140.

Arnaboldi, Valerio, Andrea Guazzini, and Andrea Passarella (2013), "Egocentric Online Social Networks: Analysis of Key Features and Prediction of Tie Strength in Facebook." *Computer Communications*, 36, 10, 1130-1144.

Barrera, Manuel, Russell E. Glasgow, H. Garth Mckay, Shawn M. Boles, and Edward G. Feil (2002), "Do Internet-Based Support Interventions Change Perceptions of Social Support?: An Experimental Trial of Approaches for Supporting Diabetes Self-Management." *American Journal of Community Psychology*, 30, 5, 637-654.

Beenen, Gerard, Kimberly Ling, Xiaoqing Wang, Klarissa Chang, Dan Frankowski, Paul Resnick, and Robert E Kraut (2004), "Using Social Psychology to Motivate Contributions to Online Communities" in *Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work,* New York: ACM, 212-221.

Belk, Russell W. (2013), "Extended Self in a Digital World." *Journal of Consumer Research*, 40, 3, 477-500.

Bickart, Barbara and Robert M. Schindler (2001), "Internet Forums as Influential Sources of Consumer Information." *Journal of Interactive Marketing*, 15, 3, 31-40.

Blomberg, Mats, Rolf Carlson, Kjell OE Elenius, Björn Granström, Joakim Gustafson, Sheri Hunnicutt, Roger Lindell, and Lennart Neovius (1993), "An Experimental Dialogue System: WAXHOLM" in *Proceedings of the Third European Conference on Speech Communication and Technology*, 1867-1870.

Bond, Robert M., Christopher J. Fariss, Jason J. Jones, Adam D. I. Kramer, Cameron Marlow, Jaime E. Settle, and James H. Fowler (2012), "A 61-Million-Person Experiment in Social Influence and Political Mobilization." *Nature*, 489, 7415, 295-298.

Booking Holdings Inc. (2020), "Kayak Chatbot." https://www.messenger.com/t/kayak. Accessed 6 July 2020.

Boshmaf, Yazan, Ildar Muslukhov, Konstantin Beznosov, and Matei Ripeanu (2013), "Design and Analysis of a Social Botnet." *Computer Networks*, 57, 2, 556-578.

Bradford, Tonya Williams, Sonya A. Grier, and Geraldine Rosa Henderson (2017), "Weight Loss through Virtual Support Communities: A Role for Identity-Based Motivation in Public Commitment." *Journal of Interactive Marketing*, 40, 9-23.

Brandtzaeg, Petter Bae and Asbjørn Følstad (2017), "Why People Use Chatbots" in *Internet Science. Insci 2017,* Cham: Springer International Publishing, 377-392.

Brindal, Emily, Jill Freyne, Ian Saunders, Shlomo Berkovsky, Greg Smith, and Manny Noakes (2012), "Features Predicting Weight Loss in Overweight or Obese Participants in a Web-Based Intervention: Randomized Trial." *Journal of Medical Internet Research*, 14, 6.

Broc & Bells (2019), "Find Your Fit. Make Healthly Connections | Broccoli & Dumbbells." https://brocnbells.com. Accessed 21 May 2019.

Broniatowski, David A, Amelia M Jamison, SiHua Qi, Lulwah AlKulaib, Tao Chen, Adrian Benton, Sandra C Quinn, and Mark Dredze (2018), "Weaponized Health Communication: Twitter Bots and Russian Trolls Amplify the Vaccine Debate." *American Journal of Public Health*, 108, 10, 1378-1384.

Brown, Jo, Amanda J Broderick, and Nick Lee (2007), "Word of Mouth Communication within Online Communities: Conceptualizing the Online Social Network." *Journal of Interactive Marketing*, 21, 3, 2-20.

Brown, John Seely and Paul Duguid (1991), "Organizational Learning and Communities-of-Practice: Toward a Unified View of Working, Learning, and Innovation." *Organization Science*, 2, 1, 40-57.

Buckle, Chase. (2016), "A Third of Online Time Spent on Social Media." https://blog.globalwebindex.net/chart-of-the-day/a-third-of-online-time-spent-on-social-media/. Accessed March 7, 2018.

Buddy System (2019), "Buddy System." https://buddysystem.eu/en. Accessed 21 May 2019.

Burke, Moira and Robert E. Kraut (2014), "Growing Closer on Facebook: Changes in Tie Strength through Social Network Site Use" in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems,* New York: ACM, 4187-4196.

Cavallo, David N., Deborah F. Tate, Amy V. Ries, Jane D. Brown, Robert F. DeVellis, and ALICE S. Ammerman (2012), "A Social Media-Based Physical Activity Intervention: A Randomized Controlled Trial." *American Journal of Preventive Medicine*, 43, 5, 527-532.

Chaniago, Muhammad Benny and Apri Junaidi (2019), "Student Presence Using RFID and" Telegram" Messenger Application: A Study in SMK Unggulan Terpadu Pgii Bandung, Indonesia." *International Journal of Higher Education*, 8, 3, 94-102.

Chawla, Nitesh V, Kevin W Bowyer, Lawrence O Hall, and W Philip Kegelmeyer (2002), "SMOTE: Synthetic Minority over-Sampling Technique." *Journal of Artificial Intelligence Research*, 16, 321-357.

Cho, Hichang, Geri Gay, Barry Davidson, and Anthony Ingraffea (2007), "Social Networks, Communication Styles, and Learning Performance in a CSCL Community." *Computers & Education*, 49, 2, 309-329.

Clement, J. (2019), "Facebook Messenger: Number of Monthly Active Users 2014-2017." https://www.statista.com/statistics/417295/facebook-messenger-monthly-active-users/. Accessed 6 July 2020.

Clementino, Davide and Lorenzo Fissore (1993), "A Man-Machine Dialogue System for Speech Access to Train Timetable Information" in *Proceedings of the Third European Conference on Speech Communication and Technology*, 1863-1866.

Cobb, Nathan K, Amanda L Graham, Beth C Bock, George Papandonatos, and David B Abrams (2005), "Initial Evaluation of a Real-World Internet Smoking Cessation System." *Nicotine & Tobacco Research*, 7, 2, 207-216.

Cobb, Nathan K., Amanda L. Graham, and David B. Abrams (2010), "Social Network Structure of a Large Online Community for Smoking Cessation." *American Journal of Public Health*, 100, 7, 1282-1289.

Cummings, Jonathon N., Brian Butler, and Robert Kraut (2002), "The Quality of Online Social Relationships." *Communications of the ACM*, 45, 7, 103-108.

Dale, Robert (2016), "The Return of the Chatbots." *Natural Language Engineering*, 22, 5, 811-817.

Davis, Clayton Allen, Onur Varol, Emilio Ferrara, Alessandro Flammini, and Filippo Menczer (2016), "Botornot: A System to Evaluate Social Bots" in *Proceedings of the 25th International Conference Companion on World Wide Web*, 273-274.

Divakaran, Pradeep Kumar Ponnamma, Adrian Palmer, Helle Alsted Søndergaard, and Roman Matkovskyy (2017), "Pre-Launch Prediction of Market Performance for Short Lifecycle Products Using Online Community Data." *Journal of Interactive Marketing*, 38, 12-28.

Du, Yong (2006), "Modeling the Behavior of Lurkers in Online Communities Using Intentional Agents" in *2006 International Conference on Computational Inteligence for Modelling Control and Automation and International Conference on Intelligent Agents Web Technologies and International Commerce (Cimca'06),* Piscataway, NJ: IEEE, 60-60.

English, Kris, Howard Goldstein, Karin Shafer, and Louise Kaczmarek (1997), "Promoting Interactions among Preschoolers with and without Disabilities: Effects of a Buddy Skills-Training Program." *Exceptional Children*, 63, 2, 229-243.

Fals-Stewart, William, Gary R. Birchler, and Michelle L. Kelley (2006), "Learning Sobriety Together: A Randomized Clinical Trial Examining Behavioral Couples Therapy with Alcoholic Female Patients." *Journal of Consulting and Clinical Psychology*, 74, 3, 579-591.

Ferrara, Emilio, Onur Varol, Clayton Davis, Filippo Menczer, and Alessandro Flammini (2016), "The Rise of Social Bots." *Communications of the ACM*, 59, 7, 96-104.

Fo, Walter SO and Clifford R O'Donnell (1975), "The Buddy System: Effect of Community Intervention on Delinquent Offenses." *Behavior Therapy*, 6, 4, 522-524.

Foster, Derek, Conor Linehan, Ben Kirman, Shaun Lawson, and Gary James (2010), "Motivating Physical Activity at Work: Using Persuasive Social Media for Competitive Step Counting" in *Proceedings of the 14th International Academic Mindtrek Conference: Envisioning Future Media Environments,* New York: ACM, 111-116.

Freyne, Jill, Shlomo Berkovsky, Stephen Kimani, Nilufar Baghaei, and Emily Brindal (2010), "Improving Health Information Access through Social Networking" in *2010 Ieee 23rd International Symposium on Computer-Based Medical Systems (CBMS),* Tharam Dillon, Daniel Rubin, William Gallagher, Amandeep Sidhu and Alexey Tsymbal eds. Piscataway, NJ: IEEE, 334-339.

Friedkin, Noah (1980), "A Test of Structural Features of Granovetter's Strength of Weak Ties Theory." *Social Networks*, 2, 4, 411-422.

Garton, Laura, Caroline Haythornthwaite, and Barry Wellman (1997), "Studying Online Social Networks." *Journal of Computer-Mediated Communication*, 3, 1, JCMC313.

Gilbert, Eric and Karrie Karahalios (2009), "Predicting Tie Strength with Social Media" in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems,* New York: ACM, 211-220.

Gilbert, Eric, Karrie Karahalios, and Christian Sandvig (2008), "The Network in the Garden: An Empirical Analysis of Social Media in Rural Life" in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems,* New York: ACM, 1603-1612.

Glass, James, Giovanni Flammia, David Goodine, Michael Phillips, Joseph Polifroni, Shinsuke Sakai, Stephanie Seneff, and Victor Zue (1995), "Multilingual Spoken-Language Understanding in the MIT Voyager System." *Speech Communication*, 17, 1-2, 1-18.

Goldstrom, Ingrid D, Jean Campbell, Joseph A Rogers, David B Lambert, Beatrice Blacklow, Marilyn J Henderson, and Ronald W Manderscheid (2006), "National Estimates for Mental Health Mutual Support Groups, Self-Help Organizations, and Consumer-Operated Services." *Administration and Policy in Mental Health and Mental Health Services Research*, 33, 1, 92-103.

Google LLC (2020), "Google Voice Actions." https://developers.google.com/voice-actions. Accessed 8 July 2020.

Graf, Bettina, Maike Krüger, Felix Müller, Alexander Ruhland, and Andrea Zech (2015), "Nombot: Simplify Food Tracking" in *Proceedings of the 14th International Conference on Mobile and Ubiquitous Multimedia,* Clemens Holzmann, René Mayrhofer, Jonna Häkkilä, Enrico Rukzio and Michael Roland eds.: Association for Computing Machinery, 360–363.

Graham, Amanda L., Kang Zhao, George D. Papandonatos, Bahar Erar, Xi Wang, Michael S. Amato, Sarah Cha, Amy M. Cohn, and Jennifer L. Pearson (2017), "A Prospective Examination of Online Social Network Dynamics and Smoking Cessation." *PloS One*, 12, 8, e0183655.

Granovetter, Mark S. (1973), "The Strength of Weak Ties." *American Journal of Sociology*, 78, 6, 1360-1380.

Gross, Doug (2011), "Apple Introduces Siri, Web Freaks Out." https://www.cnn.com/2011/10/04/tech/mobile/siri-iphone-4s-skynet/index.html. Accessed 7 July 2020.

Gruder, Charles L., Robin J. Mermelstein, Susan Kirkendol, Donald Hedeker, Siu Chi Wong, Janice Schreckengost, Richard B. Warnecke, Rebecca Burzette, and Todd Q. Miller (1993), "Effects of Social Support and Relapse Prevention Training as Adjuncts to a Televised Smoking-Cessation Intervention." *Journal of Consulting and Clinical Psychology*, 61, 1, 113-120.

Gruner, Richard L, Christian Homburg, and Bryan A Lukas (2014), "Firm-Hosted Online Brand Communities and New Product Success." *Journal of the Academy of Marketing Science*, 42, 1, 29-48.

Guhde, Jacqueline (2005), "When Orientation Ends… Supporting the New Nurse Who Is Struggling to Succeed." *Journal for Nurses in Professional Development*, 21, 4, 145-149.

Hanson, Bertil S., Sven Olof Isacsson, Lars Janzon, and Sven Eric Lindell (1990), "Social Support and Quitting Smoking for Good. Is There an Association? Results from the Population Study,"Men Born in 1914," Malmo, Sweden." *Addictive Behaviors*, 15, 3, 221-233.

Hayes, Andrew F (2018), *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach*. New York: Guilford Press.

Haythornthwaite, Caroline (2002a), "Building Social Networks Via Computer Networks: Creating and Sustaining Distributed Learning Communities" in *Building Virtual Communities: Learning and Change in Cyberspace,* K. Ann Renninger and Wesley Shumar eds. New York, NY: Cambridge University Press, 159-190.

--- (2002b), "Strong, Weak, and Latent Ties and the Impact of New Media." *The Information Society*, 18, 5, 385-401.

Hegelich, Simon and Dietmar Janetzko (2016), "Are Social Bots on Twitter Political Actors? Empirical Evidence from a Ukrainian Social Botnet" in *Proceedings of the Tenth International AAAI Conference on Web and Social Media*: The AAAI Press, 579-582.

Hektner, Joel M., Alison L. Brennan, and Gerald J. August (2017), "Incorporating Well-Adjusted Peers in a Conduct Problems Prevention Program: Evaluation of Acceptability, Fidelity, and Safety of Implementation." *School Mental Health*, 9, 1, 66-77.

Hemphill, Charles T, John J Godfrey, and George R Doddington (1990), "The ATIS Spoken Language Systems Pilot Corpus" in *Speech and Natural Language: Proceedings of a Workshop Held at Hidden Valley, Pennsylvania, June 24-27, 1990*, 96-101.

Hennrikus, Deborah, Phyllis Pirie, Wendy Hellerstedt, Harry A. Lando, Jeanne Steele, and Caroline Dunn (2010), "Increasing Support for Smoking Cessation During Pregnancy and Postpartum: Results of a Randomized Controlled Pilot Study." *Preventive Medicine*, 50, 3, 134-137.

Hoey, Louisa M, Sandra C Ieropoli, Victoria M White, and Michael Jefford (2008), "Systematic Review of Peer-Support Programs for People with Cancer." *Patient Education and Counseling*, 70, 3, 315-337.

Hoy, Matthew B (2018), "Alexa, Siri, Cortana, and More: An Introduction to Voice Assistants." *Medical Reference Services Quarterly*, 37, 1, 81-88.

Huberman, Bernardo, Daniel M. Romero, and Fang Wu (2009), "Social Networks That Matter: Twitter under the Microscope." *First Monday*, 14, 1.

Hwang, Kevin O., Allison J. Ottenbacher, Angela P. Green, M. Roseann Cannon-Diehl, Oneka Richardson, Elmer V. Bernstam, and Eric J. Thomas (2010), "Social Support in an Internet Weight Loss Community." *International Journal of Medical Informatics*, 79, 1, 5-13.

Imerman Angels (2019), "Imerman Angels | Free One-on-One Cancer Support." https://imermanangels.org. Accessed 21 May 2019.

Jeners, Nils, Petru Nicolaescu, and Wolfgang Prinz (2012), "Analyzing Tie-Strength across Different Media" in *On the Move to Meaningful Internet Systems: OTM 2012 Workshops,* P. Herrero, H. Panetto, R. Meersman and T. Dillon eds. Berlin Heidelberg: Springer, 554-563.

Jones, Jason J., Jaime E. Settle, Robert M. Bond, Christopher J. Fariss, Cameron Marlow, and James H. Fowler (2013), "Inferring Tie Strength from Online Directed Behavior." *PloS One*, 8, 1, e52168.

Kim, Hyang-Sook and S Shyam Sundar (2014), "Can Online Buddies and Bandwagon Cues Enhance User Participation in Online Health Communities?" *Computers in Human Behavior*, 37, 319-333.

Klopfenstein, Lorenz Cuno, Saverio Delpriori, Silvia Malatini, and Alessandro Bogliolo (2017), "The Rise of Bots: A Survey of Conversational Interfaces, Patterns, and Paradigms" in *Proceedings of the 2017 Conference on Designing Interactive Systems*, 555-565.

Kozinets, Robert V. (2002), "The Field Behind the Screen: Using Netnography for Marketing Research in Online Communities." *Journal of Marketing Research*, 39, 1, 61-72.

Kreijns, Karel, Paul A. Kirschner, and Wim Jochems (2003), "Identifying the Pitfalls for Social Interaction in Computer-Supported Collaborative Learning Environments: A Review of the Research." *Computers in Human Behavior*, 19, 3, 335-353.

Kukulska-Hulme, Agnes and John Pettit (2008), "Semi-Formal Learning Communities for Professional Development in Mobile Learning." *Journal of Computing in Higher Education*, 20, 2, 35-47.

Kuwata, Shigeki, Shin-ichi Taniguchi, Atsuko Kato, Kazuoki Inoue, Naoya Yamamoto, Tsuyoshi Ohkura, Kei Teramoto, Chiaki Shigemasa, and Hiroshi Kondoh (2010), "Metaboli-Net: Online Groupware System Providing Counseling Guidance for Patients with Metabolic Syndrome." *Studies in Health Technology and Informatics*, 156, 65-70.

Kviz, Frederick J., Kathleen S. Crittenden, Karen J. Madura, and Richard B. Warnecke (1994), "Use and Effectiveness of Buddy Support in a Self-Help Smoking Cessation Program." *American Journal of Health Promotion*, 8, 3, 191-201.

Lebeuf, Carlene, Margaret-Anne Storey, and Alexey Zagalsky (2017), "Software Bots." *IEEE Software*, 35, 1, 18-23.

Lee, Ran, Keon Suk Lee, Eui-Geum Oh, and Soo Hyun Kim (2013), "A Randomized Trial of Dyadic Peer Support Intervention for Newly Diagnosed Breast Cancer Patients in Korea." *Cancer Nursing*, 36, 3, E15-E22.

Levin, Esther, Shrikanth Narayanan, Roberto Pieraccini, Konstantin Biatov, Enrico Bocchieri, Giuseppe Di Fabbrizio, Wieland Eckert, Sungbok Lee, A Pokrovsky, and Mazin Rahim (2000), "The AT&T-DARPA Communicator Mixed-Initiative Spoken Dialog System" in *Proceedings of the Sixth International Conference on Spoken Language Processing*: International Speech Communication Association, 122-125.

Liberatore, Federico and L. Quijano-Sanchez (2017), "What Do We Really Need to Compute the Tie Strength? An Empirical Study Applied to Social Networks." *Computer Communications*, 110, 59-74.

Macvean, Michelle L, Victoria M White, and Rob Sanson-Fisher (2008), "One-to-One Volunteer Support Programs for People with Cancer: A Review of the Literature." *Patient Education and Counseling*, 70, 1, 10-24.

Maher, Carol A., Lucy K. Lewis, Katia Ferrar, Simon Marshall, Ilse De Bourdeaudhuij, and Corneel Vandelanotte (2014), "Are Health Behavior Change Interventions That Use Online Social Networks Effective? A Systematic Review." *Journal of Medical Internet Research*, 16, 2, e40.

Mander, Jason. (2017), "Daily Time Spent on Social Networks Rises to over 2 Hours." https://blog.globalwebindex.net/chart-of-the-day/daily-time-spent-on-social-networks/. Accessed March 7, 2018.

Mardini, Wail, Yaser Khamayseh, and Ashraf Smadi (2017), "Messenger Bot for IoT Devices" in *Proceedings of the 9th International Conference on Information Management and Engineering*: Association for Computing Machinery, 182-186.

Marsden, Peter V. and Karen E. Campbell (1984), "Measuring Tie Strength." *Social Forces*, 63, 2, 482-501.
Mathews, K. Michael, Michael C. White, Rebecca G. Long, Barlow Soper, and C. W. Von Bergen (1998), "Association of Indicators and Predictors of Tie Strength." *Psychological Reports*, 83, 3_suppl, 1459-1469.
May, Sylvia, Robert West, Peter Hajek, Andy McEwen, and Hayden McRobbie (2006a), "Randomized Controlled Trial of a Social Support ('Buddy') Intervention for Smoking Cessation." *Patient Education and Counseling*, 64, 1-3, 235-241.
--- (2006b), "Randomized Controlled Trial of a Social Support ('Buddy') Intervention for Smoking Cessation." *Patient Education and Counseling*, 64, 1, 235-241.
McCurdy, Erin E. and Christine L. Cole (2014), "Use of a Peer Support Intervention for Promoting Academic Engagement of Students with Autism in General Education Settings." *Journal of Autism and Developmental Disorders*, 44, 4, 883-893.
Merriam-Webster Inc. (2019), "Dictionary by Merriam-Webster: America's Most-Trusted Online Dictionary." https://www.merriam-webster.com. Accessed 22 May 2019.
Mesch, Gustavo and Ilan Talmud (2006), "The Quality of Online and Offline Relationships: The Role of Multiplexity and Duration of Social Relationships." *The Information Society*, 22, 3, 137-148.
Meyer, Alexandra, Adina Coroiu, and Annett Korner (2015), "One-to-One Peer Support in Cancer Care: A Review of Scholarship Published between 2007 and 2014." *European Journal of Cancer Care*, 24, 3, 299-312.
Moorhead, S. Anne, Diane E. Hazlett, Laura Harrison, Jennifer K. Carroll, Anthea Irwin, and Ciska Hoving (2013), "A New Dimension of Health Care: Systematic Review of the Uses, Benefits, and Limitations of Social Media for Health Communication." *Journal of Medical Internet Research*, 15, 4, e85.
Murray, Robert P., Janet J. Johnston, Jeffrey J. Dolce, Wondra Wong Lee, and Peggy O'Hara (1995), "Social Support for Smoking Cessation and Abstinence: The Lung Health Study." *Addictive Behaviors*, 20, 2, 159-170.
Napolitano, Melissa A., Sharon Hayes, Gary G. Bennett, Allison K. Ives, and Gary D. Foster (2013), "Using Facebook and Text Messaging to Deliver a Weight Loss Program to College Students." *Obesity*, 21, 1, 25-31.
Naylor, Rebecca Walker, Cait Poynor Lamberton, and Patricia M West (2012), "Beyond the “Like” Button: The Impact of Mere Virtual Presence on Brand Evaluations and Purchase Intentions in Social Media Settings." *Journal of Marketing*, 76, 6, 105-120.
NBA Media Ventures LLC (2020), "NBA Chatbot." https://www.messenger.com/t/nba. Accessed 6 July 2020.
Nguyen, Melanie, Yu Sun Bin, and Andrew Campbell (2012), "Comparing Online and Offline Self-Disclosure: A Systematic Review." *Cyberpsychology, Behavior, and Social Networking*, 15, 2, 103-111.
Nicholas, David B and Krista Keilty (2007), "An Evaluation of Dyadic Peer Support for Caregiving Parents of Children with Chronic Lung Disease Requiring Technology Assistance." *Social Work in Health Care*, 44, 3, 245-259.
Nonnecke, Blair and Jenny Preece (2000), "Lurker Demographics: Counting the Silent" in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems,* New York: ACM, 73-80.

O’Donnell, Clifford R and Izaak L Williams (2013), "The Buddy System: A 35-Year Follow-up of Criminal Offenses." *Clinical Psychological Science*, 1, 1, 54-66.
Obst, Patricia and Jana Stafurik (2010), "Online We Are All Able Bodied: Online Psychological Sense of Community and Social Support Found through Membership of Disability-Specific Websites Promotes Well-Being for People Living with a Physical Disability." *Journal of Community & Applied Social Psychology*, 20, 6, 525-531.
Pandorabots Inc. (2020), "Pandorabots." Accessed 6 July 2020.
Panovich, Katrina, Rob Miller, and David Karger (2012), "Tie Strength in Question & Answer on Social Network Sites" in *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work,* New York: ACM, 1057-1066.
Parkinson, Joy, Lisa Schuster, Rory Mulcahy, and Heini Maarit Taiminen (2017), "Online Support for Vulnerable Consumers: A Safe Place?" *Journal of Services Marketing*, 31, 4/5, 412-422.
Pechmann, Cornelia, Kevin Delucchi, Cynthia M. Lakon, and Judith J. Prochaska (2017), "Randomised Controlled Trial Evaluation of Tweet2Quit: A Social Network Quit-Smoking Intervention." *Tobacco Control*, 26, 2, 188-194.
Pechmann, Cornelia, Li Pan, Kevin Delucchi, Cynthia M. Lakon, and Judith J. Prochaska (2015), "Development of a Twitter-Based Intervention for Smoking Cessation That Encourages High-Quality Social Media Interactions Via Automessages." *Journal of Medical Internet Research*, 17, 2, e50.
Peckham, Jeremy (1991), "Speech Understanding and Dialogue over the Telephone: An Overview of the ESPRIT SUNDIAL Project" in *Speech and Natural Language: Proceedings of a Workshop Held at Pacific Grove, California, February 19-22, 1991*, 14-27.
Petróczi, Andrea, Tamás Nepusz, and Fülöp Bazsó (2007), "Measuring Tie-Strength in Virtual Social Networks." *Connections*, 27, 2, 39-52.
Plickert, Gabriele, Rochelle R. Cote, and Barry Wellman (2007), "It's Not Who You Know, It's How You Know Them: Who Exchanges What with Whom?" *Social Networks*, 29, 3, 405-429.
PLWC (2019), "Cancer Buddies." https://www.cancerbuddies.org.za. Accessed 21 May 2019.
Porter, Martin F. (2001), "Snowball: A Language for Stemming Algorithms." https://pdfs.semanticscholar.org/0d8f/907bb0180912d1e1df279739e45dff6853ee.pdf.
Portnoy, David B., Lori A. J. Scott-Sheldon, Blair T. Johnson, and Michael P. Carey (2008), "Computer-Delivered Interventions for Health Promotion and Behavioral Risk Reduction: A Meta-Analysis of 75 Randomized Controlled Trials, 1988--2007." *Preventive Medicine*, 47, 1, 3-16.
Powers, David MW (1998), "The Total Turing Test and the Loebner Prize" in *Proceedings of the New Methods in Language Processing and Computational Natural Language Learning*: Association for Computational Linguistics, 279-280.
Preece, Jenny, Blair Nonnecke, and Dorine Andrews (2004), "The Top Five Reasons for Lurking: Improving Community Experiences for Everyone." *Computers in human behavior*, 20, 2, 201-223.

Princeton University (2010), "About WordNet." https://wordnet.princeton.edu/. Accessed 11 July 2020.
Quijano-Sánchez, Lara, Belén Díaz-Agudo, and Juan A. Recio-García (2014), "Development of a Group Recommender Application in a Social Network." *Knowledge-Based Systems*, 71, 72-85.
Reagans, Ray (2011), "Close Encounters: Analyzing How Social Similarity and Propinquity Contribute to Strong Network Connections." *Organization Science*, 22, 4, 835-849.
Rotheram-Borus, Mary Jane, Mark Tomlinson, Margaret Gwegwe, W. Scott Comulada, Neal Kaufman, and Marion Keim (2012), "Diabetes Buddies: Peer Support through a Mobile Phone Buddy System." *The Diabetes Educator*, 38, 3, 357-365.
Seneff, Stephanie, Ed Hurley, Raymond Lau, Christine Pao, Philipp Schmid, and Victor Zue (1998), "Galaxy-II: A Reference Architecture for Conversational System Development" in *Proceedings of the Fifth International Conference on Spoken Language Processing,* Robert H. Mannell and Jordi Robert-Ribes eds.: Australian Speech Science and Technology Association, Incorporated (ASSTA), 931-934.
Setoyama, Yoko, Yoshihiko Yamazaki, and Kazuhiro Namayama (2011), "Benefits of Peer Support in Online Japanese Breast Cancer Communities: Differences between Lurkers and Posters." *Journal of Medical Internet Research*, 13, 4, e122.
Shao, Chengcheng, Giovanni Luca Ciampaglia, Onur Varol, Kai-Cheng Yang, Alessandro Flammini, and Filippo Menczer (2018), "The Spread of Low-Credibility Content by Social Bots." *Nature Communications*, 9, 1, 1-9.
Shum, Heung-Yeung, Xiao-dong He, and Di Li (2018), "From ELIZA to XiaoIce: Challenges and Opportunities with Social Chatbots." *Frontiers of Information Technology & Electronic Engineering*, 19, 1, 10-26.
Smutny, Pavel and Petra Schreiberova (2020), "Chatbots for Learning: A Review of Educational Chatbots for the Facebook Messenger." *Computers & Education*, 103862.
Snap Inc. (2019), "Snapchat - the Fastest Way to Share a Moment!" https://www.snapchat.com. Accessed 22 May 2019.
Sohrabi, Mohammad Karim and Soodeh Akbari (2016), "A Comprehensive Study on the Effects of Using Data Mining Techniques to Predict Tie Strength." *Computers in Human Behavior*, 60, 534-541.
Stella, Massimo, Emilio Ferrara, and Manlio De Domenico (2018), "Bots Increase Exposure to Negative and Inflammatory Content in Online Social Systems." *Proceedings of the National Academy of Sciences*, 115, 49, 12435-12440.
Tevyaw, Tracy O'Leary, Brian Borsari, Suzanne M. Colby, and Peter M. Monti (2007), "Peer Enhancement of a Brief Motivational Intervention with Mandated College Students." *Psychology of Addictive Behaviors*, 21, 1, 114-119.
The Monday Campaigns (2019), "Quit Monday Smoking Cessation Program." https://www.iquitmonday.org. Accessed 21 May 2019.
Trepte, Sabine, Tobias Dienlin, and Leonard Reinecke (2015), "Influence of Social Support Received in Online and Offline Contexts on Satisfaction with Social Support and Satisfaction with Life: A Longitudinal Study." *Media Psychology*, 18, 1, 74-105.
Trusov, Michael, Anand V. Bodapati, and Randolph E. Bucklin (2010), "Determining Influential Users in Internet Social Networks." *Journal of Marketing Research*, 47, 4, 643-658.
Turing, Alan M (1950), "Computing Machinery and Intelligence." *Mind*, 59, 236, 433-460.

Turner-McGrievy, Gabrielle M. and Deborah F. Tate (2013), "Weight Loss Social Support in 140 Characters or Less: Use of an Online Social Network in a Remotely Delivered Weight Loss Intervention." *Translational Behavioral Medicine*, 3, 3, 287-294.

Valle, Carmina G., Deborah F. Tate, Deborah K. Mayer, Marlyn Allicock, and Jianwen Cai (2013), "A Randomized Trial of a Facebook-Based Physical Activity Intervention for Young Adult Cancer Survivors." *Journal of Cancer Survivorship*, 7, 3, 355-368.

Wallace, Richard S (2009), "The Anatomy of ALICE" in *Parsing the Turing Test*: Springer, 181-210.

Wang, Wallace (2019), "Understanding SiriKit" in *Pro iPhone Development with Swift 5*: Springer, 491-518.

Wasko, Molly McLure and Samer Faraj (2005), "Why Should I Share? Examining Social Capital and Knowledge Contribution in Electronic Networks of Practice." *MIS Quarterly*, 29, 1, 35-57.

Weizenbaum, Joseph (1966), "ELIZA—a Computer Program for the Study of Natural Language Communication between Man and Machine." *Communications of the ACM*, 9, 1, 36-45.

West, Robert, Martin Edwards, and Peter Hajek (1998), "A Randomized Controlled Trial of a "Buddy" System to Improve Success at Giving up Smoking in General Practice." *Addiction*, 93, 7, 1007-1011.

Westmaas, J. Lee, Jeuneviette Bontemps-Jones, and Joseph E. Bauer (2010), "Social Support in Smoking Cessation: Reconciling Theory and Evidence." *Nicotine & Tobacco Research*, 12, 7, 695-707.

Wiese, Jason, Jun-Ki Min, Jason I. Hong, and John Zimmerman (2015), "You Never Call, You Never Write: Call and SMS Logs Do Not Always Indicate Tie Strength" in *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing,* New York: ACM, 765-774.

Workout Buddies (2019), "Workout Buddies App | Never Workout Alone Again." http://workoutbuddies.com. Accessed 21 May 2019.

Xiang, Rongjing, Jennifer Neville, and Monica Rogati (2010), "Modeling Relationship Strength in Online Social Networks" in *Proceedings of the 19th International Conference on World Wide Web,* New York: ACM, 981-990.

Yang, Kai-Cheng, Onur Varol, Clayton A Davis, Emilio Ferrara, Alessandro Flammini, and Filippo Menczer (2019), "Arming the Public with Artificial Intelligence to Counter Social Bots." *Human Behavior & Emerging Technologies*, 1, 1, 48-61.

Zhou, Li, Jianfeng Gao, Di Li, and Heung-Yeung Shum (2020), "The Design and Implementation of XiaoIce, an Empathetic Social Chatbot." *Computational Linguistics*, 46, 1, 53-93.

Zuyderduin, Johanna R, Valerie J Ehlers, and Dirk M Van der Wal (2008), "The Impact of a Buddy System on the Self-Care Behaviours of Women Living with HIV/AIDS in Botswana." *Health SA Gesondheid*, 13, 4, 4-15.

APPENDIX: Source Code of the Bot Intent Classifier

The Python code for the intent classifier is given below. Python is one the best and widely used programming languages for developing data products. There are lots of libraries which provide the required functionalities for developing different types of data and machine learning products from beginning to end. I also considered using a third-party framework for developing the chatbot. However, the initial results were not promising. These frameworks may work well for the type of more popular chatbots such as customer relation chatbots compared to our specific chatbot which is to work inside online support groups. On the other hand, the classifier component in those frameworks is usually a black box and developers do not have much control over it for customization. For these reasons, I developed my own code using popular classification and machine learning libraries in Python. The code is composed of the following main parts. Additional comments are added inside the code for more clarification. Comments are the lines which start with "#" sign.

Loading libraries. The first step is to load all of the libraries which provide required functionalities for operations such as data wrangling, machine learning modeling, and reporting.

Reading the data files. This is to read the raw labeled data for all groups. The data for each group is in one Excel file.

Cleaning the data. This is an important step which significantly improves the quality of data for using it in machine learning models. I cleaned both intents and messages. This includes removing links and mentions, separating emojis from the message, removing unnecessary characters such as extra white spaces, converting verbs from contracted form to uncontracted form, etc. At the end of this step, I looked for obvious inconsistencies in labeling. Inconsistencies happen when same message is labeled differently in at least two different records. I also developed functions for applying stemming and lemmatization. However, they did not improve the performance of classifier and I did not use them in the final version.

Vectorizing the text. The data cannot be used in classification models in the text format. The data must be transformed into a numerical format which can be used in machine learning models. I used tf-idf method for vectorization and I removed some stop words before vectorization.

Partitioning the data for training and test. 25% of all records were stratified randomly selected for testing. The remaining data was used for training. Since the data was totally imbalanced, I used SMOTE method for oversampling and making the number of records for triggering intents equal. The number of records for the non-triggering intent was intentionally kept higher than triggering intents. This provides more precision as compared to recall for triggering intents.

Reporting functions. In this part I define functions that later will be used for reporting the result of classification. These functions generate confusion matrix, report detailed precision, and recall for each intent, and save the results in a file in addition to show it on screen.

Training and testing. The last part uses the random forest classifier in the sklearn library for training the model. It then applied the trained model on test data and uses the earlier defined functions to report the result.

```
#Load required libraries
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from matplotlib.colors import ListedColormap
import glob
import re
import emoji
from sklearn.feature_extraction.text import TfidfVectorizer
from sklearn.model_selection import train_test_split
from imblearn.over_sampling import SMOTE
from sklearn.ensemble import RandomForestClassifier
from sklearn.metrics import classification_report, confusion_matrix, accuracy_score
import nltk
from nltk.corpus import wordnet
from nltk.stem.snowball import SnowballStemmer
from nltk.stem import WordNetLemmatizer
```

```
intents=pd.DataFrame()

#Read all data files
path="../Data/intentLabels"
all_files = glob.glob(path+"/Bot*.xlsx")
for f in all_files:
    currGroup=re.search(path+r'\\Bot(.+)_(.+).xlsx',f).group(2)
    temp=pd.read_excel(f, converters={'Tweet':str,'Correct Intent':str})
    temp.columns=["tweet","intent"]
    temp["location"]=currGroup
    intents=intents.append(temp,ignore_index=True)
```

```
#Correct coding format of intents for records
intents["intent"]=intents.intent.str.lower()\
.str.replace("'|’","")

intents["intent"]=np.where(intents.intent=="dreams","nrt_dreams",\
                np.where(intents.intent=="smokeless","smokingless",intents.intent))
```

```
#Get some descriptive statistics about the data
#Size of dataset
print(intents.shape)
#Summary of statistics
print(intents.describe())
#Frequency of intents
intents.groupby(["intent"]).size()
```

```
#Clean the tweets

#Correct some codings issues in tweets, remove links and mentions,
# and convert verbs from contracted form to uncontracted form
intents["tweet"]=intents.tweet.str.replace("â€™","'")\
.str.replace(r"http\S+","",regex=True)\
.str.replace(r"@\S+","",regex=True)\
.str.replace("’","'")\
.str.replace("n't"," not")\
.str.replace("'t"," not")\
.str.replace("'m"," am")\
.str.replace("'ll"," will")\
.str.replace("'ve"," have")\
.str.replace("'re"," are")\
.str.replace("it's","it is")

intents["tweet"] = intents["tweet"].fillna("")
```

```python
#extract all emojies from messages
dictEmoji = {':-)'  : '\U0001F642', #Smile or happy
             ':)'   : '\U0001F642',
             ':]'   : '\U0001F642',
             '=)'   : '\U0001F642',
             ':-D'  : '\U0001F603', # Big smile
             ':D'   : '\U0001F603',
             '=D'   : '\U0001F603',
             ':\'D' : '\U0001F602',#face with tears of joy
             ":')"  : '\U0001F602',
             '^_^'  : '\U0001F60A', #smiling face with smiling eyes
             '^-^'  : '\U0001F60A',
             '>:-(' : '\U0001F620',# Angry face
             '>:-o' : '\U0001F620',
             ':('   : '\U00002639', #Frowning face
             ':-('  : '\U00002639',
             ':['   : '\U00002639',
             '=('   : '\U00002639',
             ':-P'  : '\U0001F61B', #Tongue
             ':-p'  : '\U0001F61B',
             ':P'   : '\U0001F61B',
             '=p'   : '\U0001F61B',
             ':p'   : '\U0001F61B',
             ': p'  : '\U0001F61B',
             ';)'   : '\U0001F609', #wink
             ';-)'  : '\U0001F609',
             ':-*'  : '\U0001F617', #Kiss
             ':*'   : '\U0001F617',
             '<3'   : '\U00002764',#Heart
             '-_-'  : '\U0001F611',#expressionless face
             ':/'   : '\U0001F914',#Thinking
             ':-/'  : '\U0001F914',
             ':-\\' : '\U0001F914',
             ':\\'  : '\U0001F914'
            }
```

```python
#Replace text emojis like :D to equivalent binaries
def replaceTextEmojis(s):
    for k,v in dictEmoji.items():
        s=s.replace(k,v)
    return s

# The following function extracts all unicode emojis from a comment
def extract_emojis(str):
  return " ".join(c for c in str if c in emoji.UNICODE_EMOJI)
#Tranform text emojis to unicode characters
intents["tweet"]=intents.apply(lambda r:replaceTextEmojis(r.tweet),axis=1)
#separate emojies from sentence
intents["emoji"]=intents.apply(lambda r:extract_emojis(r.tweet),axis=1)
```

```python
#Only keep alphabet, numbers and spaces, transform to lower, remove extra spaces
intents["cleanTweet"]=intents.tweet.str.replace(r"[^A-Za-z0-9 ]+"," ",regex=True)\
.str.lower()\
.str.replace(r" +"," ",regex=True).str.strip() #clean whitespaces
```

```python
#Find inconsistent labels (same tweet, different labels)
uniqueCoding=intents[["cleanTweet","intent"]].drop_duplicates()
tweetCounts=uniqueCoding.groupby(["cleanTweet"]).agg({"intent":"count"})\
.rename(columns={"intent":"tweetCount"}).reset_index(drop=False)
wrongCoding=uniqueCoding.merge(tweetCounts,how="inner",on=["cleanTweet"])
wrongCoding=wrongCoding[wrongCoding.tweetCount>1].reset_index(drop=True)
wrongCoding=wrongCoding.merge(intents,how="inner",on=["cleanTweet","intent"])
wrongCoding[["cleanTweet","tweet","intent","location"]]\
.to_excel("inconsistent labeling.xlsx", index=False)
wrongCoding
```

```python
#functions for using stemmer and lemmatizer
def sentenceStemmer(s:str):
    stemmer=SnowballStemmer("english")#,ignore_stopwords=True)
    r=s.split(" ")
    for i,w in enumerate(r):
        r[i]=stemmer.stem(w)
    return " ".join(r)

tag_dict = {"J": wordnet.ADJ,
            "N": wordnet.NOUN,
            "V": wordnet.VERB,
            "R": wordnet.ADV}
wordnet_lemmatizer = WordNetLemmatizer()

def sentenceLemmatize(s:str):
    if(s==""):
        return ""
    r=s.split(" ")
    for i,w in enumerate(r):
        tag = nltk.pos_tag([w])[0][1][0].upper()
        #print(w+" "+tag+"\n")
        r[i]=wordnet_lemmatizer.lemmatize(w,pos=tag_dict.get(tag, wordnet.NOUN))
    return " ".join(r)

#intents["cleanTweet"]=intents.apply(lambda r:sentenceStemmer(r.cleanTweet),axis=1)
#intents["cleanTweet"]=intents.apply(lambda r:sentenceLemmatize(r.cleanTweet),axis=1)
```

```python
#After cleaning the data, use tf-idf to vectorize text
vectorizer=TfidfVectorizer(min_df=5,max_df=0.5,ngram_range=(1,3),\
                           stop_words=["the","a","an","to","and","as"])
features = vectorizer.fit_transform(intents["tweet"]).toarray()
print("Number of features: ",len(features[0]))
#vectorizer.get_feature_names()
```

```python
#Partition the data for test and training (stratified)
#indices_test holds the index of test records
x_train, x_test, y_train, y_test, indices_train, indices_test=train_test_split(\
    features,intents["intent"],range(len(intents.index)),test_size=.25,\
    random_state=0,stratify=intents["intent"])

print("size of train before oversampleing:",y_train.shape)
print("size of test:",y_test.shape)

#Handling imbalanced data
#use SMOTE method for oversampling the training data (except for "others")
othersIndecis=(y_train=='others')
x_trainOthers=x_train[othersIndecis]
x_trainIntents=x_train[~othersIndecis]
y_trainOthers=y_train[othersIndecis]
y_trainIntents=y_train[~othersIndecis]
oversample = SMOTE(random_state=0)
x_res, y_res = oversample.fit_resample(x_trainIntents, y_trainIntents)
x_train=np.append(x_res,x_trainOthers,axis=0)
y_train=np.append(y_res,y_trainOthers,axis=0)
print("size of train after oversampleing:",y_train.shape)
print("size of test:",y_test.shape)
#Show the size of training dataset for each intent after oversampling
pd.DataFrame(y_train).groupby(0).size()
```

```
#Function for drawing the confusion matrix
labels=["nrt_howtouse","nrt_stickissue","nrt_skinirritation","nrt_mouthirritation",\
        "nrt_nauseous","nrt_od","nrt_dreams",\
        "nrt_dontwork","nrt_itworks","nrt_runout",\
        "cravings","stress","fail","tiredness","scared",\
        "smokefree","smokingless","quitdate","support","costs",\
        "weightgain","health","cigsmell","ecigs",\
        "greetings","others"]
def drawConfusionMatrix(y_t,y_p,labels):
    cm=confusion_matrix(y_t,y_p,labels)
    pd.DataFrame(cm).to_excel("matrix.xlsx", index=False)
    lowValues = plt.cm.get_cmap('Blues', 256)
    cmap=ListedColormap(lowValues(np.linspace(0, 32, 256)))
    #I flip the confusion matrix to show the heat map correctly (the origin is lower)
    cmFlipped=np.flipud(cm)
    plt.figure(figsize=(12,12))
    plt.imshow(cmFlipped, cmap=cmap,origin="lower")#,extent=(-0.5, 2.5, 2.5, -0.5))
    tick_marks = np.arange(len(labels))
    plt.xticks(tick_marks, labels, rotation=90,fontsize=12)
    plt.yticks(list(range(len(labels)-1,-1,-1)), labels,fontsize=12)
    for i in range(len(labels)):
        for j in range(len(labels)):
            plt.text(j, i, "{:,}".format(cmFlipped[i, j]),
                     horizontalalignment="center",
                     color="white" if cmFlipped[i, j] > 20 else "black",
                     fontsize=12)
    plt.savefig('matrix.png')
    plt.show()
```

```
#Function for reporting the output
def reportClassificationResult(actual,predicted):
    drawConfusionMatrix(actual,predicted,labels)
    print(classification_report(actual,predicted,labels))
    print("Overall accuracy= ",accuracy_score(actual,predicted))
    pd.DataFrame(classification_report(actual,predicted,labels,output_dict=True))\
    .transpose().to_excel("report.xlsx")
```

```
#Use random forest for intent classification
clfRF = RandomForestClassifier(n_estimators=100,random_state=0)
clfRF.fit(x_train, y_train)
predicted=clfRF.predict(x_test)
reportClassificationResult(y_test,predicted)
```

```
#save prediction result which has actual label and predicted label
#for the test data set in an Excel file
testDataSet=intents.iloc[indices_test].reset_index(drop=True)
testDataSet["predicted"]=predicted
testDataSet.to_excel("predictionResult.xlsx", index=False)
```